

Decision making in surgical treatment of chronic low back pain : the performance of prognostic tests to select patients for lumbar spinal fusion

Citation for published version (APA):

Willems, P. C. (2011). *Decision making in surgical treatment of chronic low back pain : the performance of prognostic tests to select patients for lumbar spinal fusion*. [Doctoral Thesis, Maastricht University]. s.n.]. <https://doi.org/10.26481/dis.20111216pw>

Document status and date:

Published: 01/01/2011

DOI:

[10.26481/dis.20111216pw](https://doi.org/10.26481/dis.20111216pw)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 06 May. 2023

DECISION MAKING IN SURGICAL TREATMENT OF CHRONIC LOW BACK PAIN

*The performance of prognostic tests to select
patients for lumbar spinal fusion*

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. Mr. G.P.M.F. Mols
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 16 december 2011 om 12.00 uur

door

Paulus Cornelis Willems

Promotores:

Prof. dr. G.H.I.M. Walenkamp

Prof. dr. R.A. de Bie

Co-promotor:

Dr. M. de Kleuver, Sint Maartenskliniek Nijmegen

Beoordelingscommissie:

Prof. dr. M. van Kleeef, voorzitter

Prof. E.J. Carragee, MD, Stanford University, CA, USA

Prof. dr. W.C. Peul, Leiden Universitair Medisch Centrum

Prof. dr. R.J.E.M. Smeets

Aan mijn ouders

Voor Ankie, Anne en Thijs

Cover-Illustration and Layout:

Eric Lemmens

www.dlgraphics.nl

ISBN/EAN: 978-90-8590-051-1

The printing of this thesis was financially supported by Stichting Kliniek en Wetenschap Orthopaedie Maastricht, Dutch Spine Society, Anna Fonds/NOREF, Nederlandse Orthopaedische Vereniging, Medtronic Trading NL, Biomet NL, InSpine, Astra Tech Benelux, Baxter, Factory-CRO, Stryker NL, Synthes, Zimmer NL, Spronken orthopedie, Smeets Loopcomfort and Orthopaedie 2000

All rights reserved: No part of this publication may be reproduced or transmitted in any form without the permission of the author, or when appropriate, the publishers of the papers

CONTENTS

Chapter 1 **9**

General introduction and aims

Chapter 2 **27**

Spinal fusion for chronic low back pain:

No consensus in clinical decision making. Results of a nationwide survey among spine surgeons in the Netherlands

Revised version accepted, BMJ Open

Chapter 3a **53**

The effect of a plaster cast on lumbosacral joint motion.

An in vivo assessment with precision motion analysis system.

Spine 1997;22:1229-1234

Chapter 3b **67**

The value of a pantaloone cast test in surgical decision-making for chronic low back pain patients:

A systematic review of the literature supplemented with a prospective cohort study.

Eur Spine J 2006;15:1487-1494

Chapter 4a **89**

Provocative discography and lumbar fusion:

Is preoperative assessment of adjacent discs useful?

Spine 2007;32:1094-1099

Chapter 4b **107**

Lumbar discography: Should we use prophylactic antibiotics?

A study of 435 consecutive discograms and a systematic review of the literature.

J Spinal Disord Tech 2004;17:243-247

Contents

Chapter 5	121
Temporary external transpedicular fixation of the lumbosacral spine: A prospective, longitudinal study in 330 patients. <i>Spine 2005;30:2813-2816</i>	
Chapter 6	135
Patient selection in spinal fusion for chronic low back pain: A systematic review on the accuracy of prognostic tests. <i>Submitted</i>	
Chapter 7	159
General discussion	
Summary	173
Samenvatting	181
Dankwoord	189
Curriculum Vitae	193

CHRONIC LOW BACK PAIN

Low back pain has become a major health problem in the western world with 1-year prevalences ranging from 22 to 65% and life-time prevalences of up to 84%⁴³. About 20% of those who suffer from low back pain seek medical attention. The economic burden to society is huge and can be divided into direct costs of health care utilization (hospitalization, medication, tests and therapies) and indirect costs of lost productivity due to work absenteeism and early retirement. In a relatively small Western-European country, such as the Netherlands, the total costs of back pain in 2003 were estimated at 4.4 billion euros³⁷, which are mainly employment-related costs⁴⁰. The total annual costs of back pain in the USA exceed 100 billion dollars²⁶.

Waddell⁴¹ suggested a simple and practical classification, which divides low back pain¹ into three categories: (1) pain caused by specific spinal pathology, e.g. tumor, infection or trauma, (2) nerve root or radicular pain and (3) nonspecific low back pain, which constitutes a large heterogeneous group of patients (about 85% of total cases¹²). In these patients imaging often reveals signs of degeneration of one or more intervertebral discs, such as disc space narrowing, vertebral endplate changes^{31,45}, annular disruption^{2,22} and/or facet joint arthropathy. These degenerative findings, however, can also be observed in asymptomatic subjects²³ and thus, are nonspecific for low back pain.

The present thesis concerns patients who suffer for at least three months from nonspecific pain or discomfort in the lumbar region, with or without referred leg pain¹, and who all have signs of degeneration of the lumbar spine on plain radiographs or magnetic resonance imaging (MRI), further to be mentioned as chronic low back pain (CLBP) patients.

CLBP can result in serious physical and social restrictions and has a substantial impact on the life style of those affected. Several socioeconomic risk factors for the onset and persistence of CLBP have been identified^{4,30}: Job dissatisfaction, physically strenuous work, low education and Workers' Compensation insurance are all associated with CLBP and resulting disability²⁶. Patients out of work for 6 months due to low back pain have a 50% chance of returning to their previous job, whereas those who have been out of work for two years or more are unlikely to get

reemployed again⁴². In addition, factors such as smoking, obesity, psychological distress, depressive mood and to a lesser extent somatisation, have been reported to result in increased risk of persisting symptoms and disability³⁴. Treating this large heterogeneous group of CLBP patients with signs of degeneration of the intervertebral disc and/or facet joints on imaging, is both challenging and frustrating, as clinicians cannot rely on a true consensus among the peers for a rational therapeutic approach.

LUMBAR SPINAL FUSION FOR CLBP

The exact etiology and pathophysiology of intervertebral disc degeneration is unknown and believed to result from multiple complex interactions of biologic and biomechanical factors. Genetical predisposition appears to play an important role and several gene forms associated with disc degeneration have been identified^{3,10}. Histologically, degenerative lumbar discs reveal abnormal ingrowth of sensory nerve fibers in the endplates and the nucleus pulposus¹¹. Radial fissures in the annulus fibrosus have been linked to the ingrowth of free nerve endings (so-called nociceptors) and blood vessels with a granulation zone³³. This ingrowth of nerve endings has been suggested to correlate to the dull chronic back ache, often referred to as “discogenic pain”, that is experienced by CLBP patients, and which is exacerbated by mechanical loading of the spine. Additionally, degenerative herniated discs contain high levels of pro-inflammatory mediators and cytokines^{7,25}, such as interleukin-6, nitric oxide, prostaglandin E2²⁴ and phospholipase A2¹⁶. All these findings have strengthened the idea that disc degeneration, with or without secondary osteoarthritis of the synovial facet joints, could be a major pathway for CLBP. The assumption that the pain is largely associated with continued motion at the affected disc level has led to the concept of stabilization of a painful degenerative motion segment (i.e., intervertebral disc and facet joints) by spinal fusion to alleviate pain and reduce disability.

Spinal fusion was first described in 1889 for stabilization of vertebral segments in spinal tuberculosis²⁰ and has a well-established role in the treatment of spinal fractures and deformities. Its role in the treatment of CLBP, however, remains controversial^{28,38}. Epidemiological research reveals large variation in the amount of

spinal fusion operations between countries⁹ and even between different regions within the same country up to 20-fold ranges in fusion rate have been reported^{27,44}. Although socio-economic factors and the availability of spine surgery in a region or country may differ, such wide geographic variations suggest a poor level of professional consensus on the indications for spinal fusion in CLBP patients.

In a Cochrane review in 1999¹⁹, no evidence was found on the effectiveness of fusion for lumbar degenerative disc disease (DDD) or low back pain, as compared to natural history, placebo, or conservative treatment. Despite the absence of concurrent reports of clarified indications or improved efficacy, there was a 220% increase in the rate of lumbar spine fusion surgery from 1990 to 2001 in the USA (Figure 1)¹³. This rise

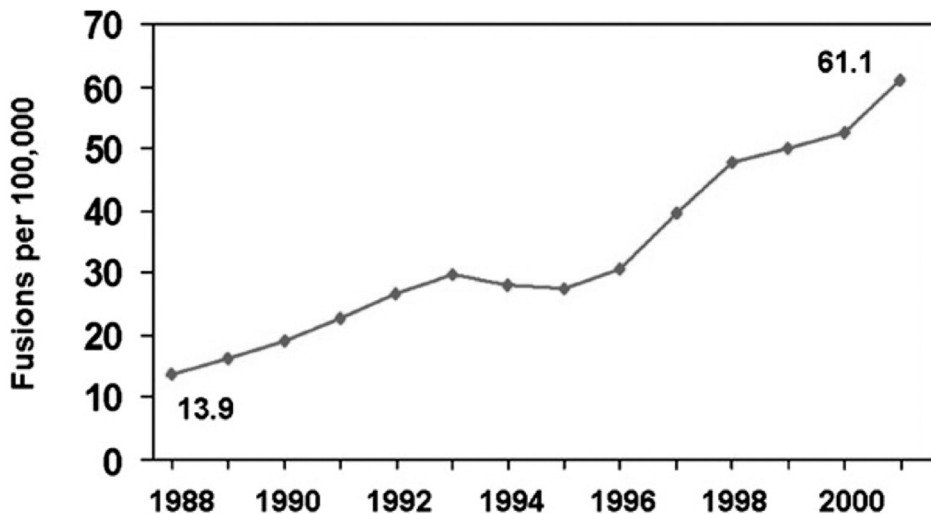


Figure 1. The rise of spinal fusion procedures for degenerative conditions in the USA between 1988 and 2001. Note the rapid increase after 1996, when fusion cages were approved. (Courtesy from Mr Richard A Deyo, *J Am Board Fam Med* 2009;22:62-68, reproduced with permission).

coincided with the promotion and FDA-approval of the fusion cages (interbody implants to enhance stability and fusion) in 1996. As a result, Medicare expenses for lumbar fusion increased from 75 million dollars in 1992 to 482 million dollars in 2003, which was about half of total USA spending for spine surgery⁴⁴.

In 2001, a Swedish randomized controlled trial (RCT)¹⁸ did show a better outcome in patients treated with spinal fusion compared to patients who received standard conservative care, although at longer follow-up this beneficial effect attenuated¹⁷. Two more recent RCTs, which compared fusion surgery to cognitive behavioural based exercise therapy⁶ or an intensive rehabilitation program¹⁵, showed similar results for surgery and nonsurgery at 1 year, and at 2 years follow-up, respectively. In the latter trial, fusion had a higher complication rate⁴⁶ and appeared to be less cost-effective than intensive rehabilitation³⁶. From these trials one might conclude that a considerable amount of patients do not improve after surgery and that spinal fusion is a questionable treatment for the entire heterogeneous group of CLBP patients. In a recently published survey²¹, renowned spine surgeons did not consider complications or bad surgery, but bad patient selection, wrong diagnosis and disproportionate expectations of patients to be the major factors for poor outcome in spinal surgery. Therefore, it would be of great value to know whether the results of fusion surgery can be improved by the identification of subgroups of CLBP patients who truly benefit from spinal fusion.

At present, all routine diagnostic tools to select the right patient for lumbar fusion appear to be inadequate. History taking and physical examination may reveal psychological distress or social factors which can lead to chronicity of pain and need treatment^{14,32}. There are, however, no specific physical findings to predict the outcome of lumbar fusion⁸. The association of CLBP with findings on imaging is weak³⁹. Degenerative signs on plain radiographs do not appear to be correlated with persisting symptoms of CLBP²⁹ and magnetic resonance imaging (MRI), which has been recommended as the imaging study of choice for the initial evaluation of patients with CLBP³⁵, shows high false-positive rates of degenerative findings in asymptomatic people, especially in the elderly^{5,23}.

PROGNOSTIC TESTS TO AID IN CLINICAL DECISION MAKING

To identify those patients who will benefit from fusion surgery, many surgeons rely on tests that are assumed to predict the outcome of spinal fusion. In order to improve the results of lumbar fusion for CLBP, and aiming for consensus on the indications for fusion, it is essential to know the role and value of these prognostic tests in clinical practice.

In the present thesis, we will focus on the most commonly used prognostic tests:

(1) Immobilization in a thoracolumbosacral orthosis (TLSO)

In order to mimic the immobilizing effect of a lumbar fusion, a standard manufactured brace or corset can be prescribed or a plaster cast applied on the standing patient usually from just below the shoulder blades to the middle of the sacrum (Figure 2). For immobilization of level L5-S1, one hip is fixed in 10° of flexion with the thigh-piece ending 10 cm proximal to the patella. Patients are expected to wear the TLSO or cast continuously for at least 2 to 4 weeks and are encouraged to perform as much daily life activities as possible. In case of significant pain relief while wearing the TLSO, the test is considered positive.



Figure 2. Immobilization test by a thoracolumbosacral orthosis, e.g., a plaster cast. Extension of the cast with a thigh piece is optional and assumed to better immobilize the lumbosacral joint.

(2) Identification of painful segments by provocative discography

This procedure can be performed on an outpatient basis. Under sterile conditions and with the patient in lateral decubitus position, a stilette needle is advanced into the intended disc space (Figure 3). When correct placement of the needle tip in the center of the disc is verified by biplanar fluoroscopy, a contrast agent is injected under pressure. If this injection provokes pain similar to the patient's usual pain, (i.e., concordant pain) and, preferably, if 1 or 2 discs adjacent to the suspect disc do not elicit concordant pain on provocation (i.e., control levels), the test can be considered positive. In addition, plain radiographs or CT-scans can be obtained to determine the extent of degeneration of the injected disc(s).

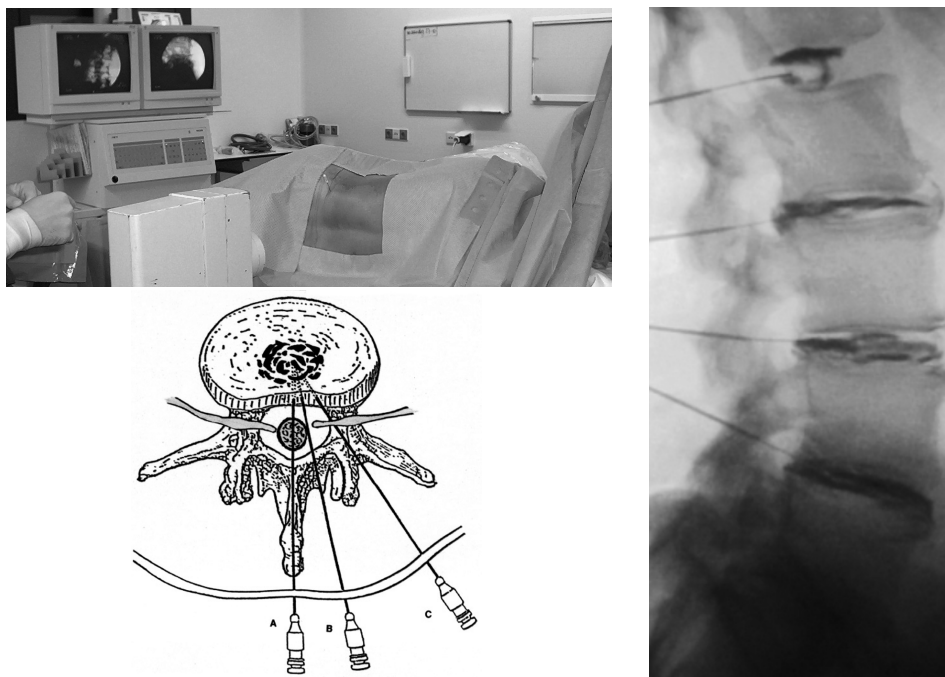


Figure 3. Provocative discography is performed under sterile conditions and with biplanar fluoroscopic control. The intended disc space is injected with contrast agent to provoke usual back pain. In addition, the amount of degeneration can be visualized on radiographs or CT.

(3) Trial immobilization by temporary external transpedicular fixation (TETF)

For this test, patients are hospitalized for 7 to 10 days. Under general anesthesia, antibiotic prophylaxis and fluoroscopic control, two screws are inserted percutaneously through the pedicles into the vertebra above, and two screws into the vertebra below the suspect disc(s), respectively. Postoperatively, the protruding screw ends are fixed externally with two vertical bars, which immobilizes the disc(s) of interest, thus mimicking a lumbar fusion (Figure 4). In case of adequate pain relief, the test is considered positive. Optionally, immobilization can be discontinued - without knowledge of the patient - by fixing the bars horizontally, which, of course, should annul pain relief (i.e., placebo test).

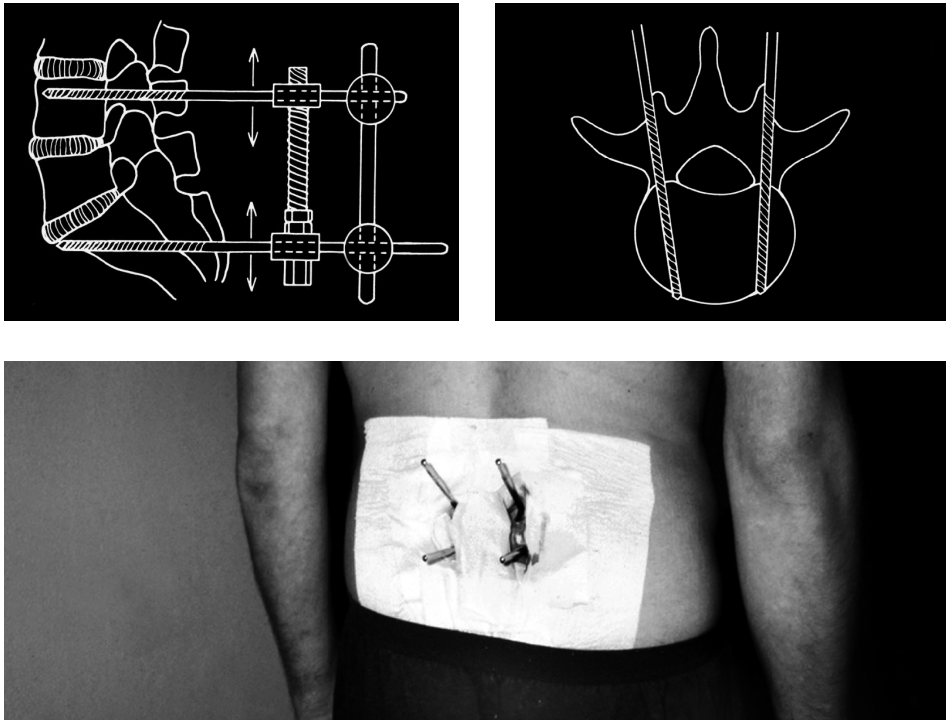


Figure 4. Temporary external transpedicular fixation trial: By externally fixing two vertebrae, one or more vertebral motion segments are immobilized, thus simulating the effect of a spinal fusion.

AIMS AND OUTLINE OF THIS THESIS

The aims of the present thesis are to assess the value of three commonly used prognostic tests (i.e., immobilization by a TLSO, provocative discography, trial immobilization by TETF) for the clinical outcome of lumbar spinal fusion in CLBP patients, and to evaluate the use and appreciation of these tests by spine surgeons in clinical practice.

Aim 1: Is there consensus among spine surgeons regarding the use of prognostic tests for lumbar spinal fusion in CLBP patients?

In chapter 2 the results of a nationwide survey among spine surgeons in the Netherlands are presented. The surgeons are questioned about their opinion on prognostic factors and the use and appreciation of prognostic tests for lumbar spinal fusion in CLBP patients.

Aim 2: Does immobilization by a pantaloone cast truly minimize lumbosacral motion?

In chapter 3a the hypothesized working mechanism of a pantaloone cast, i.e. mimicking a lumbar fusion by minimization of lumbosacral mobility, is studied. In patients, admitted for a temporary external transpedicular fixation test (TETF), infrared light markers are fixed to the protruding pins of two spinal levels. In this way, three-dimensional motion between these levels can be analyzed optoelectronically during dynamic test conditions. Measurements are performed before and after the application of a pantaloone cast.

Aim 3: Can immobilization by a thoracolumbosacral orthosis (TLSO) predict the clinical outcome of spinal fusion for CLBP?

In chapter 3b a systematic review of the literature supplemented with the results of a prospective patient cohort study is presented in order to assess the value of a TLSO in predicting the long-term clinical outcome of fusion for CLBP.

Aim 4: Does provocative discography of adjacent segments predict the long-term clinical outcome of spinal fusion for CLBP?

In chapter 4a the results of provocative discography of levels adjacent to an intended lumbar fusion are presented in a cohort of patients for whom the decision to perform fusion has been based on a TETF trial. The relation between preoperative discography results and long-term clinical outcome is assessed.

Aim 5: What is the incidence of postdiscography discitis and is there a need for routine antibiotic prophylaxis?

Although low in incidence, intervertebral discitis is the most feared complication of provocative discography. In chapter 4b the risk of postdiscography discitis and the need for routine prophylactic antibiotics are studied in a cohort of 200 consecutive patients, and by means of a systematic review of the literature.

Aim 6: Can temporary external transpedicular fixation (TETF) help to predict the clinical outcome of spinal fusion in CLBP patients with an equivocal indication for surgery?

In chapter 5 the middle- and long-term results of TETF as a test to predict the clinical outcome of lumbar fusion, are presented in a group of CLBP patients with a doubtful indication for surgery. Suspect lumbar motion segments are temporarily stabilized by external fixation. In case of adequate pain relief, lumbar fusion may be indicated. The test includes a placebo trial, in which the patients are unaware whether the lumbar segmental levels are fixed or dynamized.

Aim 7: What is the prognostic accuracy of the most commonly used tests in clinical practice to predict the outcome of spinal fusion for CLBP?

Chapter 6 presents the results of a systematic literature review concerning the prognostic accuracy of tests that are currently used in clinical practice to predict the outcome of lumbar spinal fusion for CLBP. The tests of interest are magnetic resonance imaging (MRI), immobilization by a thoracolumbosacral orthosis (TLSO), TETF, provocative discography and facet joint infiltration.

In chapter 7, the general discussion focuses on the present and possible future role of lumbar spinal fusion as a treatment option for CLBP. The so-called biopsychosocial model for the evaluation of CLBP patients is explained. Triage in a multidisciplinary setting or Spine Centre is advocated, and the importance of treatment in a stepwise fashion within the current, widely supported active approach of CLBP, is emphasized.

REFERENCES

1. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006;15 Suppl 2:S192-300.
2. Aprill C, Bogduk N. High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. *Br J Radiol* 1992;65:361-9.
3. Battie MC, Videman T, Levalahti E, et al. Genetic and environmental effects on disc degeneration by phenotype and spinal level: a multivariate twin study. *Spine* 2008;33:2801-8.
4. Bigos SJ, Battie MC, Fisher LD. Methodology for evaluating predictive factors for the report of back injury. *Spine (Phila Pa 1976)* 1991;16:669-70.
5. Boden SD, McCowin PR, Davis DO, et al. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:1178-84.
6. Brox JI, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913-21.
7. Burke JG, Watson RW, McCormack D, et al. Intervertebral discs which cause low back pain secrete high levels of proinflammatory mediators. *J Bone Joint Surg Br* 2002;84:196-201.
8. Carragee EJ. Clinical practice. Persistent low back pain. *N Engl J Med* 2005;352:1891-8.
9. Cherkin DC, Deyo RA, Loeser JD, et al. An international comparison of back surgery rates. *Spine* 1994;19:1201-6.
10. Cheung KM, Chan D, Karppinen J, et al. Association of the Taq I allele in vitamin D receptor with degenerative disc disease and disc bulge in a Chinese population. *Spine* 2006;31:1143-8.
11. Coppes MH, Marani E, Thomeer RT, et al. Innervation of "painful" lumbar discs. *Spine (Phila Pa 1976)* 1997;22:2342-9; discussion 9-50.
12. Deyo RA, Bass JE, Walsh NE, et al. Prognostic variability among chronic pain patients: implications for study design, interpretation, and reporting. *Arch Phys Med Rehabil* 1988;69:174-8.
13. Deyo RA, Gray DT, Kreuter W, et al. United States trends in lumbar fusion surgery for degenerative conditions. *Spine (Phila Pa 1976)* 2005;30:1441-5; discussion 6-7.
14. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *Jama* 1992;268:760-5.

15. Fairbank J, Frost H, Wilson-MacDonald J, et al. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *Bmj* 2005;330:1233.
16. Franson RC, Saal JS, Saal JA. Human disc phospholipase A2 is inflammatory. *Spine (Phila Pa 1976)* 1992;17:S129-32.
17. Fritzell P, Hagg O, Jonsson D, et al. Cost-effectiveness of lumbar fusion and nonsurgical treatment for chronic low back pain in the Swedish Lumbar Spine Study: a multicenter, randomized, controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2004;29:421-34; discussion Z3.
18. Fritzell P, Hagg O, Wessberg P, et al. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001;26:2521-32; discussion 32-4.
19. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 1999;24:1820-32.
20. Hadra BE. Wiring of the spinous process in injury and Pott's disease. *Trans Am Orthop Assoc* 1891;4:206-10.
21. Haefeli M, Elfering A, Aebi M, et al. What comprises a good outcome in spinal surgery? A preliminary survey among spine surgeons of the SSE and European spine patients. *Eur Spine J* 2008;17:104-16.
22. Ito M, Incorvaia KM, Yu SF, et al. Predictive signs of discogenic lumbar pain on magnetic resonance imaging with discography correlation. *Spine (Phila Pa 1976)* 1998;23:1252-8; discussion 9-60.
23. Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 1994;331:69-73.
24. Kang JD, Georgescu HI, McIntyre-Larkin L, et al. Herniated lumbar intervertebral discs spontaneously produce matrix metalloproteinases, nitric oxide, interleukin-6, and prostaglandin E2. *Spine (Phila Pa 1976)* 1996;21:271-7.
25. Kang JD, Stefanovic-Racic M, McIntyre LA, et al. Toward a biochemical understanding of human intervertebral disc degeneration and herniation. Contributions of nitric oxide, interleukins, prostaglandin E2, and matrix metalloproteinases. *Spine* 1997;22:1065-73.
26. Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *J Bone Joint Surg Am* 2006;88 Suppl 2:21-4.

27. Katz JN. Lumbar spinal fusion. Surgical rates, costs, and complications. *Spine* 1995;20:78S-83S.
28. Krismer M. Fusion of the lumbar spine. A consideration of the indications. *J Bone Joint Surg Br* 2002;84:783-94.
29. Kumar MN, Jacquot F, Hall H. Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease. *Eur Spine J* 2001;10:309-13.
30. McIntosh G, Frank J, Hogg-Johnson S, et al. Prognostic factors for time receiving workers' compensation benefits in a cohort of patients with low back pain. *Spine (Phila Pa 1976)* 2000;25:147-57.
31. Modic MT, Masaryk TJ, Ross JS, et al. Imaging of degenerative disk disease. *Radiology* 1988;168:177-86.
32. Nelson MA, Allen P, Clamp SE, et al. Reliability and reproducibility of clinical findings in low-back pain. *Spine (Phila Pa 1976)* 1979;4:97-101.
33. Peng B, Wu W, Hou S, et al. The pathogenesis of discogenic low back pain. *J Bone Joint Surg Br* 2005;87:62-7.
34. Pincus T, Burton AK, Vogel S, et al. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 2002;27:E109-20.
35. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 2: assessment of functional outcome. *J Neurosurg Spine* 2005;2:639-46.
36. Rivero-Arias O, Campbell H, Gray A, et al. Surgical stabilisation of the spine compared with a programme of intensive rehabilitation for the management of patients with chronic low back pain: cost utility analysis based on a randomised controlled trial. *Bmj* 2005;330:1239.
37. Slobbe LCJ KG, Smit JM, Groen J, Meerding WJ, Polder JJ. Kosten van ziekten in Nederland 2003. RIVM rapport 270751010. Bilthoven: RIVM, 2006.
38. Turner JA, Ersek M, Herron L, et al. Patient outcomes after lumbar spinal fusions. *Jama* 1992;268:907-11.

39. van Tulder MW, Assendelft WJ, Koes BW, et al. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine* 1997;22:427-34.
40. van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. *Pain* 1995;62:233-40.
41. Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine (Phila Pa 1976)* 1987;12:632-44.
42. Waddell G. Low back pain: a twentieth century health care enigma. *Spine* 1996;21:2820-5.
43. Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *J Spinal Disord* 2000;13:205-17.
44. Weinstein JN, Lurie JD, Olson PR, et al. United States' trends and regional variations in lumbar spine surgery: 1992-2003. *Spine* 2006;31:2707-14.
45. Weishaupt D, Zanetti M, Hodler J, et al. Painful Lumbar Disk Derangement: Relevance of Endplate Abnormalities at MR Imaging. *Radiology* 2001;218:420-7.
46. Wilson-MacDonald J, Fairbank J, Frost H, et al. The MRC spine stabilization trial: surgical methods, outcomes, costs, and complications of surgical stabilization. *Spine* 2008;33:2334-40.

2

Spinal fusion for chronic low back pain: No consensus in clinical decision making Results of a nationwide survey among spine surgeons in the Netherlands

Willems PC, de Bie RA, Öner FC, Castelein RM, de Kleuver M

Revised version accepted, BMJ Open

ABSTRACT

Objectives. To assess the appreciation and use of prognostic patient factors and predictive tests for clinical decision making in spinal fusion for chronic low back pain, and to relate surgeons' opinion in practice to findings from the literature.

Design and setting. Nationwide survey among spine surgeons.

Participants. Surgeon members of the Dutch Spine Society were questioned on their treatment strategy for patients with chronic low back pain.

Main outcome measures. The surgeons' opinion on prognostic patient factors known from the literature (e.g., work status, litigation and psychosocial status), and their use of predictive tests (e.g., trial immobilization, provocative discography) for patient selection, were addressed. In addition, the influence of surgeon specific factors (age, discipline, clinical experience) on decision making was assessed.

Results. The comments from 62 surgeons (response rate of 70%) were analyzed. Forty-four surgeons (71%) had a clinical experience of 10 years or more. There was a statistically significant lack of uniformity of opinion in 7 of the 11 items on prognostic factors and 8 of the 11 items on predictive tests, respectively. Apart from the use of discography and long multisegment fusions, differences in training or clinical experience did not appear to be of significant influence on treatment strategy.

Conclusion. The present survey consistently showed a lack of consensus among Dutch spine surgeons in the management of CLBP. There was a surprisingly large variation in decision making for such a relatively small, highly urbanized country as the Netherlands. Despite high levels of training and continuous medical education, decision making for fusion surgery to treat CLBP does not appear to have a uniform evidence base in daily clinical practice.

INTRODUCTION

Chronic low back pain (CLBP) has become one of the main causes of disability in the industrialized world with reported life-time prevalences of up to 85%³³. In the Netherlands, a small Western European country (16.5 million inhabitants) with a relatively high rate of spine surgery⁸, the total annual costs of back pain in 2003 were estimated at 4.4 billion euros (approximately 6 billion US dollars), which are mainly employment-related costs (lost productivity due to work absenteeism)²⁹. The total costs of back pain in the USA exceed 100 billion dollars per year²² and Medicare expenses for lumbar fusion alone have increased from 75 million dollars in 1992 to 482 million dollars in 2003, which is about half of total spending for spine surgery³⁴.

Treating CLBP patients is both challenging and frustrating because physicians cannot rely on a true consensus among the peers for a rational therapeutic approach. Although spinal fusion of a painful or degenerative segment can be beneficial to some patients, it remains a controversial treatment for CLBP^{24,30}. In the first Cochrane review in 1999, no evidence on the effectiveness of fusion for lumbar degenerative disc disease (DDD) or low back pain was found as compared to natural history, placebo or conservative treatment¹⁸. In the updated Cochrane review in 2005¹⁹, two randomized controlled trials (RCT) on fusion for CLBP were included. First, a Swedish trial reported a better outcome of patients treated with spinal fusion compared to patients who received standard conservative care¹⁷, although at longer follow-up this beneficial effect attenuated¹⁶. Next, a Norwegian RCT that compared fusion surgery to cognitive behavioural based exercise therapy⁵ showed similar results for both treatment modalities at 1 year follow-up. Similarly, in the more recent British spine stabilization trial no clear evidence was found that spinal fusion was more beneficial than an intensive rehabilitation program at 2 years follow-up¹⁵. Moreover, fusion had a much higher complication rate in this trial and appeared to be less cost-effective than intensive rehabilitation^{27,37}.

Undoubtedly, spinal fusion truly helps some patients with CLBP. If these patients could be identified in advance, a subgroup of patients with a successful outcome after fusion could be selected. Although in the literature several prognostic factors and predictive tests for surgical outcome have been reported,

there are no generally acknowledged clinical guidelines to determine which patient may benefit from fusion.

Epidemiological research reveals large variation in the amount of spinal fusion operations between countries and even up to 20-fold ranges in fusion surgery rate between different regions within the same country^{23,34}. Although socio-economic factors and the availability of spine surgery in a region or country may differ, such wide geographic variations suggest a poor level of professional consensus. Understanding contributory factors in the decision making process of surgeons, may clarify some of these observed variations.

The aim of this study was to conduct a national survey among Dutch spine surgeons in order to gain insight in their treatment strategy for CLBP. Specifically, the use of predictive tests for spinal fusion surgery and the surgeons' opinion on prognostic patient factors were assessed and related to findings from the literature. Finally, the degree of consensus in clinical decision making was determined.

MATERIALS AND METHODS

A 25-question survey (see Appendix) was sent by mail to all surgeon members of the Dutch Spine Society, by Memic, a Center for Data and Information Management, University of Maastricht, the Netherlands (www.memic.unimaas.nl). In an accompanying letter the background rationale for the enquiry, as well as the voluntary and confidential nature was stressed and the surgeons were reassured that individual comments would remain anonymous.

The questionnaire concerned the selection for spinal fusion of patients with low back pain caused by degenerative lumbar disc disease without signs of neurological deficit, spinal stenosis, deformity or spondylolisthesis and in the absence of trauma, tumor or infections. This group was further referred to as chronic low back pain (CLBP) patients. For clarity the questionnaire had first been evaluated and revised by a clinical researcher and two orthopaedic surgeons. Most questions could be answered according to a 5-point Likert scale. Surgeon specific factors (e.g., discipline, clinical experience), the influence of patient factors (prognostic factors as reported in literature), and the use of predictive tests (e.g., provocative discography) for patient selection, were addressed. The respondents

were specifically asked to rely on their own individual opinion and management in practice.

After a second call by mail to those who had not responded, the final inclusion date for analysis was set on April 1, 2009. Data were entered into Excel™ (Microsoft, Corp., Redmond, WA) and all inconsistencies were resolved. Unanswered questions were coded as missing. Descriptive statistics was used in which all frequencies were based on the number of valid responders.

For analysis the answers on the 5-point Likert scale were merged into one intermediate option ("neutral") and 2 opposite categories ("always/almost always" versus "never/almost never" and "fully/globally agree" versus "globally/fully disagree"). The data were processed with Statistical Package for the Social Sciences software (SPSS, Inc., Chicago, IL). Pearson's chi-square test was used to evaluate whether surgeon specific factors were associated with clinical decision making. Uniformity of opinion was defined to be present if 70% or more of the respondents answered similarly. In other words, there was no consensus if the proportion of the largest category was statistically significantly lower than 70% (Pearson's chi-square test). Differences in mean values rating the impact of factors on decision making, were tested by Independent t-test for equality of means. The level of significance was set at $p=0.05$.

RESULTS

Nine of the 150 surveyed surgeons (89 orthopaedic surgeons and 61 neurosurgeons) had ended their professional career and 9 respondents stated not to perform spinal surgery anymore. Of the remaining 132 active spine surgeons, 93 (70%) completed and returned the questionnaire. Thirty-one of the 93 respondents (33%) declared not to perform spinal fusion for CLBP and were excluded from further analysis. The characteristics of the final group of 62 respondents are listed in Table 1. The level of experience for neurosurgeons and orthopaedic surgeons was equal: 11 of 16 (69%) versus 33 of 46 (72%) worked 10 years or more in clinical practice, respectively.

Table 1. Characteristics of the 62 respondents

	Orthopaedic surgeons (n)	Neurosurgeons(n)	All respondents (n)
No. of respondents	46	16	62
Age			
< 50 years	22	10	32
≥ 50 years	24	6	30
Clinical experience			
< 10 years	13	5	18
≥ 10 years	33	11	44
Type of hospital			
University/specialized	13	5	18
General	33	11	44
No. of fusions for CLBP/year			
1-10	24	9	33
10-25	9	6	15
25-50	7	1	8
≥ 50	6	0	6

Abbreviation: CLBP = chronic low back pain; n = number

Prognostic factors

The respondents' comments on prognostic factors are listed in Table 2. For 7 of the 11 items there was no consensus (significantly less than 70% uniformity of opinion).

More than 70% of the respondents would fuse patients over 60 years old for back pain. This was related to caseload: Only 4 of 29 surgeons (14%) who performed more than 10 fusions for CLBP annually, did not operate the elderly, versus 13 of 33 surgeons (39%) with a lower caseload ($p=0.024$). Years of clinical experience or specialty did not appear to be of influence ($p=0.504$, and $p=0.690$, respectively).

Only 1 of 15 neurosurgeons fused patients below 20 for back pain, versus 14 of 46 orthopaedic surgeons ($p=0.063$).

Eighteen orthopaedic surgeons performed fusion of 3 or more levels for CLBP, whereas no neurosurgeon would ($p=0.003$).

Table 2. Respondents' opinion to what extent patient-specific prognostic factors influence their clinical decision making in the treatment of CLBP. The numbers listed are percentages of valid responses

Maximum number of levels for fusion	1 level 18 (30.5)	2 levels 23 (39.0)	3 or more levels 18 (30.5)	p-value* p<0.001
Minimum age patient	Under 20 yrs 15 (24.6)	20 to 30 yrs 25 (41.0)	30 yrs or more 21 (34.4)	p<0.001
Maximum age patient	40 to 50 yrs 5 (8.1)	50 to 60 yrs 12 (19.4)	60 yrs or more 45 (72.5)	NS
Minimal length conservative therapy	Less than 6 mo 3 (4.8)	6 mo to 1 yr 36 (58.1)	1 yr or more 23 (37.1)	NS
Maximum Body Mass Index (BMI)	Under 31 29 (46.8)	31 to 37 18 (29.0)	37 or more 15 (24.2)	p<0.001
Maximum number of cigarettes / day	0 29 (47.5)	1 to 20 7 (11.4)	20 or more 25 (40.9)	p<0.001
Referral overweight patients to dietician	Always 29 (46.8)	Sometimes 20 (32.3)	Never 13 (21.0)	p<0.001
Psychological screening referral	Always 10 (16.2)	Sometimes 28 (45.2)	Never 24 (38.7)	p<0.001
Different criteria for primary DDD versus prior spine surgery	Agree 44 (71.0)	Neutral 8 (12.9)	Disagree 10 (16.1)	NS
Work status affects outcome	Agree 29 (46.7)	Neutral 17 (27.4)	Disagree 16 (25.9)	p<0.001
Litigation procedures affect outcome	Agree 43 (69.3)	Neutral 9 (14.5)	Disagree 10 (16.2)	NS

Abbreviation: DDD = degenerative disc disease, NS = not significant.

*Chi-square test: p<0.05 means significantly less than 70% consensus, NS implies uniformity.

Predictive tests

The surgeons' appreciation and use of predictive tests are listed in Tables 3a and 3b, respectively. Apart from MRI, there was no uniformity regarding the value of these tests for clinical decision making.

Table 3a. Respondents' opinion on predictive tests for clinical decision making. The numbers listed are valid responses and respective percentages

Predictive test	Agree (%)	Neutral (%)	Disagree (%)	p-value*
MRI sufficient for decision making	10 (16.1)	11 (17.7)	41 (66.1)	NS
Cast immobilization valuable test	25 (40.3)	15 (24.2)	22 (35.5)	<0.001
Cast immobilization too unpleasant	11 (17.7)	16 (25.8)	35 (56.5)	0.028
PD proven valuable test	23 (37.7)	16 (26.2)	22 (36.0)	<0.001
PD too many complications	3 (4.9)	14 (23.0)	44 (72.1)	NS
TETF valuable test	8 (13.4)	33 (55.0)	19 (31.6)	0.011
TETF too many complications	20 (32.7)	31 (50.8)	10 (16.4)	0.001

Abbreviations: MRI = magnetic resonance imaging, PD = provocative discography, TETF = temporary external transpedicular fixation, NS = not significant.

*Chi-square test: $p < 0.05$ means significantly less than 70% consensus, NS implies uniformity.

Table 3b. The use of predictive tests by the surgeons in clinical practice. The numbers listed are valid responses and their respective percentages

Use of test	Always (%)	Sometimes (%)	Never (%)	p-value*
Facet joint blocks	5 (8.1)	32 (51.6)	25 (40.3)	0.002
Cast immobilization	20 (32.8)	23 (37.7)	18 (29.6)	<0.001
PD	25 (42.4)	10 (16.9)	24 (40.7)	<0.001
TETF	0 (0.0)	3 (4.9)	58 (95.1)	NS

Abbreviations : PD = provocative discography, TETF = temporary external transpedicular fixation, NS = not significant.

*Chi-square test: $p < 0.05$ means significantly less than 70% consensus, NS implies uniformity.

Mainly orthopaedic surgeons (21 of 46, versus 2 of 16 neurosurgeons, $p=0.025$) considered provocative discography to be a valid predictor of fusion. Spine surgeons working in general hospitals (20 of 43), appeared to believe more in the test than academic surgeons did (3 of 18, $p=0.028$). There was no relation with clinical experience ($p=0.406$).

Patient history and imaging were valued significantly higher for clinical decision making than were predictive tests, psychological screening or patient preferences (all respective comparisons: $p < 0.01$, Independent t-test, Figure 1).

In the evaluation of CLBP no other predictive tests than those mentioned in Tables 3a and 3b were used on a regular basis.

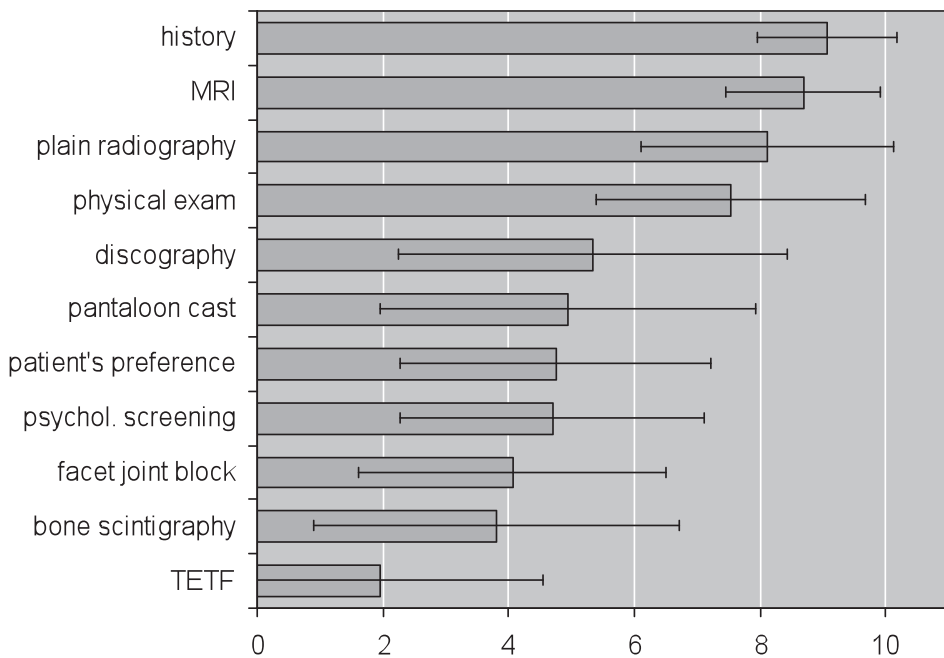


Figure 1. The importance of listed factors in clinical decision making (presented as mean \pm sd) as rated by the respondents on a scale from 0 (no importance) to 10 (maximal importance). Abbreviations: MRI = magnetic resonance imaging, TETF = temporary external transpedicular fixation.

Individual decision making in clinical practice

Experience was rated highest (mean \pm sd, 8.0 ± 1.7) as factor of influence on clinical decision making, as compared to findings from literature (7.7 ± 1.1 , $p = 0.26$), scientific courses (7.3 ± 1.4 , $p = 0.01$), and training (6.8 ± 2.8 , $p < 0.01$) (Figure 2).

Twenty-seven (45%) surgeons responded to have a protocol for decision

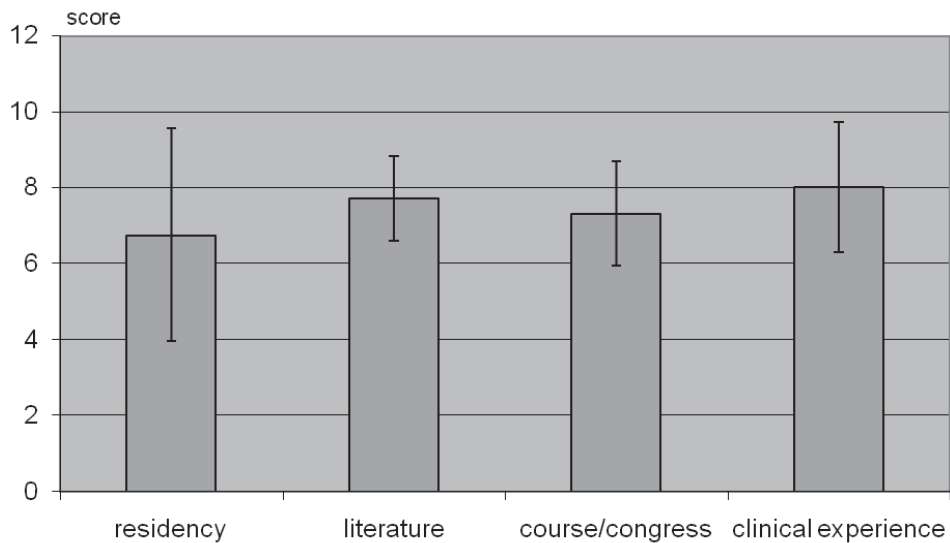


Figure 2. Factors that influence clinical decision making for CLBP (presented as mean \pm sd), as rated by the respondents on a scale from 0 (no influence) to 10 (maximal influence).

making to which they frequently or always adhered. Of those 35 respondents who did not have such a protocol, 23 (68%) replied that there should be guidelines. In other words, 50 respondents (83%) felt that clinical guidelines in the management of CLBP patients are prerequisite.

DISCUSSION

This study presents the results of the first nationwide survey among spine surgeons regarding clinical decision making for spinal fusion in CLBP patients. The response rate was adequate (70%) and the majority of the respondents (71%) had extensive clinical experience in spinal surgery (Table 1). A considerable heterogeneity in the use and appreciation of predictive tests was observed. Prognostic patient factors were not consistently incorporated in clinical decision making.

Comparison with related research

Almost three quarters (73%) of the respondents fused patients above 60 for CLBP,

whereas older age is an acknowledged predictor of poor outcome¹⁰.

Despite the fact that 2 or 3-level fusions have been reported to have proven higher rates of pseudarthrosis with lower patient satisfaction as compared to single level fusions^{10,30}, over 30% of the respondents would consider fusion of 3 levels or more.

Although fusion surgery for CLBP is not recommended unless 2 years of conservative treatment have failed¹, 63% of the surgeons felt that less than 1 year of conservative therapy is enough to consider fusion.

Obesity is an independent risk factor for CLBP and surgery in these patients is significantly associated with major complications, such as thrombo-embolism and infection²⁵. Nevertheless, 53% of the respondents would operate for CLBP on obese patients (BMI > 31) and 24% on the morbid obese (BMI > 37). Less than half of the surgeons (47%) consistently referred overweight patients to a dietician.

There was no consensus regarding smoking, which is known to be an independent risk factor for CLBP¹¹ and associated with worse results of spinal fusion³⁷. About 41% would fuse heavy smokers, whereas 48% would not operate smokers for CLBP.

Psychologically stressful work has been associated with LBP and disability²² and it has been reported that psychological distress, depressive mood and somatisation lead to an increased risk of chronicity²⁶. In addition, presurgical depression is associated with worse patient outcome after lumbar fusion¹⁰. Only 16% of the respondents referred CLBP patients routinely for psychological screening and 39% never referred for this purpose at all.

There is strong evidence that clinical interventions are not effective in returning CLBP patients back to work once they have been off work for a longer time¹. About half of the respondents agreed that the work status of LBP patients affects outcome considerably and 69% acknowledged that litigation or workers' compensation are of great influence on decision making in CLBP as they have been associated with persisting LBP and disability²².

Two-thirds (66%) of the respondents considered findings on plain radiographs and MRI-scan alone to be insufficient for surgical decision making in CLBP (Table 3a). This is in accordance with the literature indicating that degenerative or black discs on MRI do not appear to have a strong clinical relevance^{4,21} and that there is

no correlation between radiographic signs of degeneration and clinical outcome³¹.

Opinion differed about trial immobilization with a pantaloone cast: 40% of the respondents agreed that it is a valuable test and 36% disagreed. This resembles conflicting reports from the literature claiming that the test is not predictive of fusion outcome² or that only in highly selected patient groups the pantaloone cast test may be of value³⁵.

Provocative discography is a controversial test, which is highly variable in chronic pain patients and can also be positive in pain-free individuals⁷. Its value in predicting the outcome of fusion for CLBP is debated^{9,36}, which was reflected in the completely contradictory respondents' opinions. Trial immobilization with a temporary external fixator (TETF) is known for its high complication rate³ and because of ambiguous results, its use is not recommended¹³. In the present survey TETF was not frequently used (94% never used it) and only 13% believed in its proven predictive value.

Lumbar facet injections have been reported not to be predictive of either arthrodesis or nonsurgical treatment of CLBP¹⁴. Accordingly, only 8% used facet joint blocks on a regular basis as a predictor of spinal fusion.

Strengths and limitations

This survey focused on surgeon members of the Dutch Spine Society whose practice may not reflect that of all surgeons performing lumbar fusion for CLBP. This may have produced a selection bias. It is reasonable, however, to expect that surgeons with a special interest in the spine are exactly those to be most aware of guidelines and research findings in the field.

To define consensus we chose for uniformity of opinion of 70% or more of the respondents. We felt that this level of agreement should be sufficient for implementation in guidelines. Such a cut-off level remains, of course, arbitrary and debatable.

The introduction of an interviewer bias could be avoided by employing Memic, Center for Data and Information Management, as a neutral intermediary. In this way, surgeons could feel free to answer what they personally felt or practiced, as opposed to what they thought would be considered "correct".

For statistical analysis the 5-point Likert scale responses were merged into 3

categories, which may have simplified the respondents' opinion on the management of CLBP in practice.

Clinical relevance and implications for daily practice

The lack of consensus among spine surgeons as found in the present survey could not be explained by differences in training or clinical experience. Apart from the use of discography and long multilevel fusions, the surgeons' discipline and years in practice did not appear to be of significant influence on treatment strategy. More likely, the observed heterogeneity of opinion reflects the absence of consistent high quality evidence for the validity of prognostic factors, predictive tests and lumbar fusion itself for CLBP. As there is no generally acknowledged superior approach, substantial variations that exist between practices are caused by clinical uncertainty as to what constitutes the best of care for these patients.

In a recently published survey²⁰ expert spine surgeons considered bad patient selection and disproportionate preoperative expectations to be the major factors for poor outcome in spinal surgery. To provide a reliable estimation of the effectiveness of surgery, assessment of the individual patient is needed. CLBP is not a diagnosis but rather a symptom in patients with varying stages of impairment and disability. Psychological and social factors may affect outcome and should be known³². Good patient counselling should be evidence based, i.e., determined by the individual surgeon's expertise and expectation of treatment success⁶, combined with the best available clinical evidence from systematic research²⁸. At present, consistent evidence on tests or tools that reliably predict the outcome of fusion is lacking¹², which hampers the implementation of consensus guidelines. As long as we cannot reliably identify a subgroup of patients who will benefit from fusion, the results of fusion surgery will remain similar to those of regimented rehabilitation programmes, but with more costs and the risk of substantial complications.

Research should focus on identifying those patients for whom fusion surgery provides a better outcome than conservative care. Improved patient selection may establish a role for lumbar fusion as the treatment of choice for a subgroup of CLBP patients, and would enable the installment of guidelines for surgical decision making. Such guidelines are needed, not only for patient counseling, but also for communication with insurers, policy makers and other health care providers.

Conclusion

The present survey consistently showed a lack of consensus among spine surgeons in the management of CLBP. There was a surprisingly large variation in surgical decision making for such a relatively small, highly urbanized country as the Netherlands. Despite high levels of training and continuous medical education, patient selection for fusion surgery in the treatment of CLBP does not appear to have a uniform evidence base in clinical practice.

Acknowledgments

The authors would like to thank the Dutch Spine Society for financial support, and acknowledge its members for their participation in this study.

REFERENCES

1. Airaksinen O, Brox JJ, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006;15 Suppl 2:S192-300.
2. Axelsson P, Johnsson R, Stromqvist B, et al. Orthosis as prognostic instrument in lumbar fusion: no predictive value in 50 cases followed prospectively. *J Spinal Disord* 1995;8:284-8.
3. Bednar DA. Failure of external spinal skeletal fixation to improve predictability of lumbar arthrodesis. *J Bone Joint Surg Am* 2001;83-A:1656-9.
4. Boden SD, McCowin PR, Davis DO, et al. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:1178-84.
5. Brox JJ, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913-21.
6. Campbell C, Guy A. 'Why can't they do anything for a simple back problem?' A qualitative examination of expectations for low back pain treatment and outcome. *J Health Psychol* 2007;12:641-52.
7. Carragee EJ, Tanner CM, Khurana S, et al. The rates of false-positive lumbar discography in select patients without low back symptoms. *Spine* 2000;25:1373-80; discussion 81.
8. Cherkin DC, Deyo RA, Loeser JD, et al. An international comparison of back surgery rates. *Spine* 1994;19:1201-6.
9. Colhoun E, McCall IW, Williams L, et al. Provocation discography as a guide to planning operations on the spine. *J Bone Joint Surg Br* 1988;70:267-71.
10. DeBerard MS, Masters KS, Colledge AL, et al. Outcomes of posterolateral lumbar fusion in Utah patients receiving workers' compensation: a retrospective cohort study. *Spine* 2001;26:738-46; discussion 47.
11. Deyo RA, Bass JE. Lifestyle and low-back pain. The influence of smoking and obesity. *Spine* 1989;14:501-6.
12. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery - the case for restraint. *N Engl J Med* 2004;350:722-6.
13. Elmans L, Willems PC, Anderson PG, et al. Temporary external transpedicular fixation of the lumbosacral spine: a prospective, longitudinal study in 330 patients. *Spine* 2005;30:2813-6.

14. Esses SI, Moro JK. The value of facet joint blocks in patient selection for lumbar fusion. *Spine* 1993;18:185-90.
15. Fairbank J, Frost H, Wilson-MacDonald J, et al. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *Bmj* 2005;330:1233.
16. Fritzell P, Hagg O, Jonsson D, et al. Cost-effectiveness of lumbar fusion and nonsurgical treatment for chronic low back pain in the Swedish Lumbar Spine Study: a multicenter, randomized, controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2004;29:421-34; discussion Z3.
17. Fritzell P, Hagg O, Wessberg P, et al. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001;26:2521-32; discussion 32-4.
18. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 1999;24:1820-32.
19. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis: updated Cochrane Review. *Spine* 2005;30:2312-20.
20. Haefeli M, Elfering A, Aebi M, et al. What comprises a good outcome in spinal surgery? A preliminary survey among spine surgeons of the SSE and European spine patients. *Eur Spine J* 2008;17:104-16.
21. Jarvik JJ, Hollingworth W, Heagerty P, et al. The Longitudinal Assessment of Imaging and Disability of the Back (LAIDBack) Study: baseline data. *Spine* 2001;26:1158-66.
22. Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *J Bone Joint Surg Am* 2006;88 Suppl 2:21-4.
23. Katz JN. Lumbar spinal fusion. Surgical rates, costs, and complications. *Spine* 1995;20:785-835.
24. Krismer M. Fusion of the lumbar spine. A consideration of the indications. *J Bone Joint Surg Br* 2002;84:783-94.
25. Patel N, Bagan B, Vadera S, et al. Obesity and spine surgery: relation to perioperative complications. *J Neurosurg Spine* 2007;6:291-7.
26. Pincus T, Burton AK, Vogel S, et al. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 2002;27:E109-20.

27. Rivero-Arias O, Campbell H, Gray A, et al. Surgical stabilisation of the spine compared with a programme of intensive rehabilitation for the management of patients with chronic low back pain: cost utility analysis based on a randomised controlled trial. *Bmj* 2005;330:1239.
28. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *Bmj* 1996;312:71-2.
29. Slobbe LCJ KG, Smit JM, Groen J, Meerding WJ, Polder JJ. Kosten van ziekten in Nederland 2003. RIVM rapport 270751010. Bilthoven: RIVM, 2006.
30. Turner JA, Ersek M, Herron L, et al. Patient outcomes after lumbar spinal fusions. *Jama* 1992;268:907-11.
31. van Tulder MW, Assendelft WJ, Koes BW, et al. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine* 1997;22:427-34.
32. Waddell G. Low back pain: a twentieth century health care enigma. *Spine* 1996;21:2820-5.
33. Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *J Spinal Disord* 2000;13:205-17.
34. Weinstein JN, Lurie JD, Olson PR, et al. United States' trends and regional variations in lumbar spine surgery: 1992-2003. *Spine* 2006;31:2707-14.
35. Willems PC, Elmans L, Anderson PG, et al. The value of a pantaloon cast test in surgical decision making for chronic low back pain patients: a systematic review of the literature supplemented with a prospective cohort study. *Eur Spine J* 2006;15:1487-94.
36. Willems PC, Elmans L, Anderson PG, et al. Provocative discography and lumbar fusion: is preoperative assessment of adjacent discs useful? *Spine (Phila Pa 1976)* 2007;32:1094-9; discussion 100.
37. Wilson-MacDonald J, Fairbank J, Frost H, et al. The MRC spine stabilization trial: surgical methods, outcomes, costs, and complications of surgical stabilization. *Spine* 2008;33:2334-40.

APPENDIX

Questionnaire on decision making for lumbar spinal fusion in chronic low back pain patients

- | | | |
|---|--|---|
| 1 | What is your discipline? | 1 Neurosurgery
2 Orthopaedic surgery
3 Other,
..... |
| 2 | What is your age? | 1 Under 30 years
2 30 to 40 years
3 40 to 50 years
4 50 to 60 years
5 60 years or older |
| 3 | Since when do you perform spinal surgery? | 1 Less than 1 year
2 1 to 5 years
3 5 to 10 years
4 10 to 15 years
5 15 years or more |
| 4 | In what kind of hospital do you work?
(more than one answer possible) | 1 University hospital
2 General teaching hospital
3 General nonteaching hospital
4 Specialized hospital
5 Other,
..... |

The next questions concern the indication for lumbar spinal fusion (or lumbar total disc replacement if appropriate) in patients with low back pain caused by degenerative lumbar disc disease without signs of neurological deficit, spinal stenosis, deformity or spondylolisthesis and in the absence of trauma, tumor, infections or other consuming illnesses, further to be referred to as chronic low back pain (CLBP) patients

- | | | |
|----|--|---|
| 5 | How many lumbar fusions do you perform each year in CLBP patients? | 1 0
2 1 to 10
3 10 to 25
4 25 to 50
5 50 or more |
| | | |
| 6 | How many total disc replacements do you perform each year in CLBP patients? | 1 0
2 1 to 10
3 10 to 25
4 25 to 50
5 50 or more |
| | | |
| 7 | What is for you the maximum number of levels to be fused in CLBP patients? | 1 1
2 2
3 3
4 4 or more
5 No maximum |
| | | |
| 8a | What is for you the absolute minimum age of a CLBP patient to be considered for lumbar fusion? | 1 Under 20 years
2 20 to 30 years
3 30 to 40 years
4 40 years or more
5 No minimum age |

- | | | |
|-----|--|---|
| 8b | What would be for you the absolute maximum age of a CLBP patient to be considered for lumbar fusion? | <ul style="list-style-type: none">1 Under 40 years2 40 to 50 years3 50 to 60 years4 60 years or older5 No maximum age |
| 9 | How long should a CLBP patient at least have followed conservative therapy in order to be considered for lumbar fusion? | <ul style="list-style-type: none">1 0 to 6 months2 6 months to 1 year3 1 to 2 years4 2 years or longer5 No minimum |
| 10a | What would be for you the maximum weight of a 1.80 meter long male CLBP patient in order to be considered for lumbar fusion? | <ul style="list-style-type: none">1 Less than 80 kg2 80 to 100 kg3 100 to 120 kg4 120 kg or more5 No maximum weight |
| 10b | Do you send overweight CLBP patients to a dietician before considering lumbar fusion? | <ul style="list-style-type: none">1 Always2 Frequently3 Sometimes4 Seldom5 Never |
| 11 | What is for you the maximum number of cigarettes a CLBP patient is allowed to smoke in order to be considered for lumbar fusion? | <ul style="list-style-type: none">1 0 cigarettes per day2 1 to 10 cigarettes per day3 10 to 20 cigarettes per day4 20 or more cigarettes per day5 No maximum |

- | | |
|--|---|
| 12 Do you send CLBP patients for psychological screening before considering lumbar fusion? | 1 Always
2 Frequently
3 Sometimes
4 Seldom
5 Never |
|--|---|

You are requested to indicate whether you agree or not with the following statements.

- | | Fully
agree | Partially
agree | Neutral | Partially
disagree | Fully
disagree |
|---|----------------|--------------------|----------|-----------------------|-------------------|
| 13 The preoperative selection criteria for CLBP patients who had spine surgery before are substantially different from those for CLBP patients without prior spine surgery. | 1 | 2 | 3 | 4 | 5 |
| 14 The work status (Full or partial disability, long term sick leave) of a CLBP patient is of great influence on your decision to perform lumbar fusion. | 1 | 2 | 3 | 4 | 5 |
| 15 Involvement in litigation or workers compensation processes is of great influence on your decision making. | 1 | 2 | 3 | 4 | 5 |

Chapter 2

- | | | | | | | |
|----|--|----------|----------|----------|----------|----------|
| 16 | Plain radiographs and MRI-findings in CLBP patients are sufficient for your decision to perform lumbar fusion. | 1 | 2 | 3 | 4 | 5 |
|----|--|----------|----------|----------|----------|----------|

The next statements and questions concern clinical tests that may be helpful in decision making for lumbar fusion in CLBP patients.

- | | | | | | | |
|-----|---|----------|----------|----------|----------|----------|
| 17a | Trial immobilization in a plaster jacket or pantaloons cast is a proven valuable test for decision making in CLBP patients. | 1 | 2 | 3 | 4 | 5 |
|-----|---|----------|----------|----------|----------|----------|

- | | | | | | | |
|-----|--|----------|------------|---------------|----------|----------|
| | | Always | Frequently | Some
times | Seldom | Never |
| 17b | Do you use this trial immobilization in a cast in CLBP patients? | 1 | 2 | 3 | 4 | 5 |

- | | | | | | | |
|-----|--|----------------|--------------------|----------|-----------------------|-------------------|
| | | Fully
agree | Partially
agree | Neutral | Partially
disagree | Fully
disagree |
| 17c | Trial immobilization in a cast is too unpleasant for the patient to be executed. | 1 | 2 | 3 | 4 | 5 |

- | | | | | | | |
|-----|---|----------------|--------------------|----------|-----------------------|-------------------|
| | | Fully
agree | Partially
agree | Neutral | Partially
disagree | Fully
disagree |
| 18a | Provocative discography is a proven valuable test for decision making in CLBP patients. | 1 | 2 | 3 | 4 | 5 |

Survey: Clinical decision making for spinal fusion

	Always	Frequently	Some times	Seldom	Never
18b Are CLBP patients in your practice selected for fusion by provocative discography?	1	2	3	4	5
18c Provocative discography has too many complications to be executed.	Fully agree 1	Partially agree 2	Neutral 3	Partially disagree 4	Fully disagree 5
19a Temporary external transpedicular fixation (TETF) of one or more segments is a proven valuable for decision making in CLBP patients.	Fully agree 1	Partially agree 2	Neutral 3	Partially disagree 4	Fully disagree 5
19b Do you use TETF as a tool for decision making in CLBP patients?	Always 1	Frequently 2	Some times 3	Seldom 4	Never 5

Chapter 2

	Fully agree	Partially agree	Neutral	Partially disagree	Fully disagree
19c TETF has too many complications to be executed in CLBP patients.	1	2	3	4	5

	Always	Frequently	Some times	Seldom	Never
20 Are CLBP patients in your practice selected for fusion by facet joint blocks?	1	2	3	4	5

21 Do you use other tests as a selective tool for lumbar fusion in CLBP patients?	1	No
	2	Yes,
	
	

Could you rate on a scale ranging from 0 (no importance) to 10 (maximum importance) how important you consider each of the following items as a selective tool for lumbar fusion in CLBP patients

	0 to 10
22a Plain radiographs
22b MRI-scan
22c Bone scintigraphy
22d History
22e Physical examination
22f Psychological screening
22g Patient's preferences
22h Facet joint blocks
22i Trial immobilization by pantaloon cast
22j Lumbar provocative discography
22k Temporary external transpedicular fixation

Could you rate on a scale ranging from 0 (no influence) to 10 (maximal influence) to what extent your policy regarding the operative management of CLBP patients has been influenced by the following factors:

	0 to 10
23a Knowledge acquired during residency / training
23b Knowledge from the literature
23c Knowledge from courses or congresses
23d Knowledge based on clinical impression and experience

	Very satisfied	Fairly satisfied	Neutral	Fairly unsatisfied	Very unsatisfied
	1	2	3	4	5
24 Are you satisfied with the results of the management of CLBP patients in your practice?					

25a Are there protocols or guidelines in your clinic as to what CLBP patients can be considered for lumbar fusion?	1	Yes, please continue with question 25b
	2	No, please continue with question 25c

	Always	Frequently	Sometimes	Seldom	Never
	1	2	3	4	5
25b If yes, do you adhere to these guidelines for every CLBP patient in your practice?					

	Fully agree	Partially agree	Neutral	Partially disagree	Fully disagree
	1	2	3	4	5
25c If no, do you think there should be guidelines for the management of CLBP patients?					

3a

The effect of a plaster cast on lumbosacral joint motion

An in vivo assessment with precision motion analysis system

Willems PC, Nienhuis B, Sietsma M, van der Schaaf DB, Pavlov PW

Spine 1997;22:1229-1234

ABSTRACT

Study Design. This study was conducted to assess the effect of a plaster cast on the mobility of the lumbosacral joint in 10 patients with chronic low back pain. During static and dynamic exercises, movements between the proximal vertebra (L4 or L5) and the sacrum were registered in 10 patients without a support and after the application of a plaster cast, and with and without unilateral hip immobilization, respectively.

Objectives. To investigate whether plaster casts actually immobilize the lumbosacral joint.

Summary of Background Data. The presumed stabilizing effect of a lumbar orthosis on the lumbosacral joint has been the subject of many studies in the past years, and contradictory reports have been published.

Methods. The measurements were performed by means of Precision Motion Analysis System, an optoelectronic three-dimensional motion analysis system using infrared light. The patients were asked to perform maximal spinal flexion to extension, maximal pelvic tilt (static test conditions), and to walk within the measurement volume (dynamic test condition). This procedure was repeated with the patients wearing a plaster cast with and without unilateral hip fixation. Mobility was expressed in translations and rotations around three axes. For statistical analysis, repeated measurements two-way analysis of variance was used.

Results. Considerable rotations were found only in the sagittal plane. Both plaster casts appeared to decrease mobility during the static test conditions. During the dynamic test condition, however, no significant decrease of mobility of the lumbosacral joint by either of the casts could be observed. In both cast conditions, considerably more sagittal rotation was found during walking than with the other two exercises.

Conclusion. In the sagittal plane, a plaster cast with or without unilateral hip immobilization can decrease motion during spinal flexion-extension. This stabilizing effect on the lumbosacral joint could not be observed during walking.

INTRODUCTION

External lumbar supports are widely used in the management of low back pain. In addition to conservative treatment, they also are applied to ascertain the effect of immobilization before surgery, when spinal fusion is being considered, and to support the spine after surgery until fusion has occurred. The most important effect expected from the support is the restriction of lumbosacral movements¹⁹. Varied and unpredictable effects, however, with even an increase in lumbosacral movements, have been observed^{11,14}. Adding unilateral hip immobilization to the lumbar support should provide maximal stabilization of the lumbosacral joint, although this has been contradicted by a recent study². Most studies reported in the literature measured lumbosacral displacements in static conditions⁷. The purpose of the present study was to assess the effect of a plaster cast with and without unilateral hip immobilization on lumbosacral movements in static and dynamic test conditions.

METHODS

Patients

Ten consecutive patients (7 men, 3 women; mean age, 37 years; range, 25-49 years) with intractable low back pain volunteered for this study. All patients had suspected instability of the lower lumbar spine and had been admitted for a temporary, percutaneous, transpedicular fixation test^{6,12,15}.

Two Schanz screws (5 mm diameter) were inserted transpedicularly into the fourth (four patients) or fifth (six patients) lumbar vertebra, and two screws were inserted into the sacrum (Table 1). By external fixation of the screw ends, the effect of immobilization of the lumbosacral spine can be analyzed. The present experiments were conducted before the external fixator was applied, thus permitting free mobility of the vertebrae involved.

Material

The plaster casts were applied on the standing patient and extended posteriorly from just below the shoulder blades to the middle of the sacrum, anteriorly from the xiphoid process to just above the os pubis, and laterally from around the lower ribs to

3 cm below the level of the anterior superior iliac spines. For unilateral hip immobilization, the left hip was fixed in 10° of flexion with the thighpiece ending 10 cm proximal to the patella. In both casts, a hole was created with a margin of approximately 2 cm around the protruding screw ends to obtain free range of motion (Figure 1A). For each patient, the two casts were applied by the same orthotist.

Procedure

Measurements were performed with Precision Motion Analysis System (PRIMAS, University of Technology, Delft, The Netherlands), an optoelectronic measuring system using infrared light^{8,9}. Two video cameras were placed 4 m behind the patient under an angle of 80 degrees, thus creating an effective measurement volume of approximately 3 m in length, 1.5 m in width, and 1.5 m in height (Figure 2). Two frames, each containing three reflective spherical markers with a diameter of 3 cm, were attached to the blunt outer end of two transpedicular screws: one to a proximal screw and one to a sacral screw (Figure 1B). Before each procedure, the patient was placed in a standardized position in the field of measurement, and the orientation of the frame attached to the sacrum was determined. Thus, relative to the sacrum, a three-dimensional local coordinate system (x-, y-, and z-axis) was defined that coincided with the frontal, horizontal, and sagittal planes for each patient. Motion of the proximal frame relative to the sacral frame then was expressed in translations (accuracy ± 1 mm) along the three axes and rotations (± 0.3 degrees, accuracy for static and moving markers) around the three axes⁸. Continuous registrations with a sampling frequency of 100 Hz were acquired with the patient in a neutral erect position, maximal spinal flexion, maximal spinal extension, maximal pelvic tilt, and while walking. All exercises were performed three times, and the patients were encouraged to perform the exercises to the limit of their endurance. The same sequence was repeated with the patient wearing a plaster cast with and without unilateral hip fixation, respectively (Figure 1C). In this way every patient acted as his own control.

Data Analysis

For all exercises, the median values of rotation and translation were determined. Next, in all three planes, the ranges of rotation and translation from spinal flexion

to extension and during pelvic tilt were calculated. For each step of the walking exercise, the ranges between the maxima and minima were determined. In each patient, the difference (delta) from the ranges of the condition without support to the ranges obtained in both casts (*i.e.*, range in cast without hip fixation minus range without support and range in cast with hip fixation minus range without support) were calculated. These deltas were compared by means of a repeated measurements two-way analysis of variance (ANOVA, SPSS for Windows, Release Version 6.1, SPSS Inc., Chicago, IL), taking a probability of 5% as the level of significance.

RESULTS

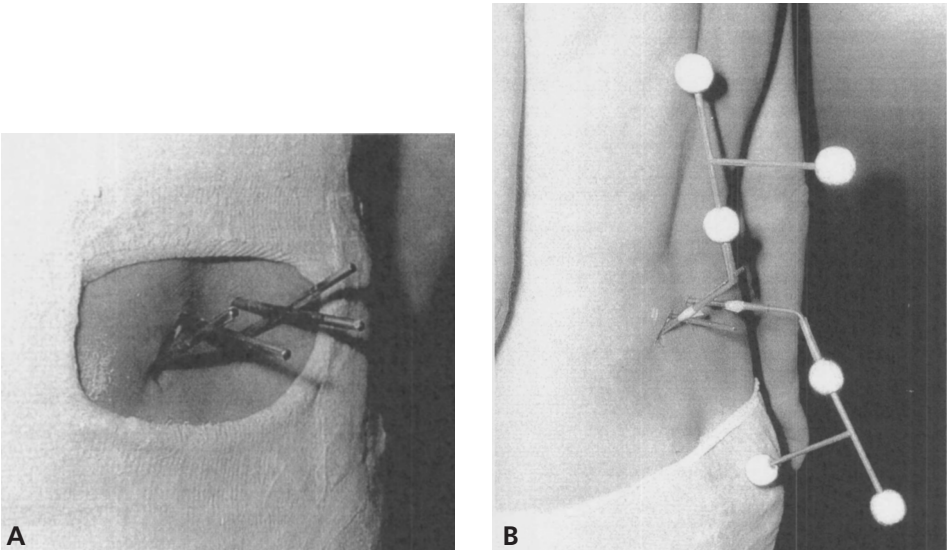
Translations were of the order of 1 mm, which was considered too small compared with the accuracy of the system, and were excluded from further analysis.⁸ Noteworthy rotations (more than 1.0 degree) were only found in the sagittal plane. The ranges in sagittal rotation during the three sequences are presented in Table 1. In the condition without support, as well as in both casts, interindividual differences were considerable, and a stabilizing effect was not predictable for any individual patient.

In the flexion-extension exercise (Figure 3A), the patients with motion measured across two segments (L4-S1, numbers 1,2 and 3) showed considerably higher sagittal rotation with significant reduction (high-negative deltas) by both plaster casts than the other patients ($F(1,8) = 6.1$; $P = 0.039$). In the pelvic tilt exercise (Figure 3B), initial sagittal rotations were very low, so negative deltas hardly could be seen. In the walking exercise (Figure 3C), deltas in many patients were practically zero, and statistical analysis for all patients showed no significant difference in sagittal rotation between the condition without support and either of the plaster casts (statistical power = 0.8 for detecting a true difference of 1 degree in rotation = $1.25 \times \text{SD}$, given $\alpha = 0.05$)¹³.

In both cast conditions, considerably more sagittal rotation was found during walking than during the other two exercises, reaching almost the level of significance ($F(1,9) = 4.99$; $P = 0.052$).

Table 1. Patient Characteristics and Angles of Lumbosacral Sagittal Rotation (°) in the Unrestricted Condition and in Both Plaster Cast Conditions, While Performing Spinal Flexion-Extension, Pelvic Tilt, and Walking

				Spinal Flexion-Extension		
Patient No.	Sex	Age (yr)	Fixation Level	No Plaster Cast	Cast - Hip Flexion	Cast + Hip Flexion
1	M	43	L4-S1	3.0	4.9	0.6
2	M	46	L4-S1	6.7	2.4	2.2
3	M	41	L4-S1	6.3	1.7	1.3
4	M	30	L4-L5	0.5	1.9	2.0
5	F	34	L5-S1	0.9	0.5	0.5
6	M	49	L5-S1	2.8	0.1	0.2
7	F	25	L5-S1	0.1	0.3	0.6
8	M	37	L5-S1	0.7	0.4	1.1
9	F	28	L5-S1	1.1	0.7	1.2
10	M	34	L5-S1	2.0	2.3	0.8



	Pelvic Tilt			Walking		
	No Plaster	Cast - Hip	Cast + Hip	No Plaster	Cast - Hip	Cast + Hip
	Cast	Fixation	Fixation	Cast	Fixation	Fixation
	3.5	3.4	0.2	2.2	1.0	2.6
	0.2	0.3	1.9	1.2	2.8	1.8
	0.5	1.2	0.6	2.1	2.1	2.0
	0.5	2.7	1.9	1.1	0.9	1.1
	2.1	0.9	0.9	1.6	1.7	1.4
	0.4	1.6	0.7	2.4	2.7	1.7
	0.7	1.0	0.8	1.0	1.0	0.8
	1.2	1.5	1.2	0.8	6.4	2.1
	1.2	1.6	1.4	3.2	1.1	3.5
	0.7	0.4	0.7	2.5	3.7	1.9

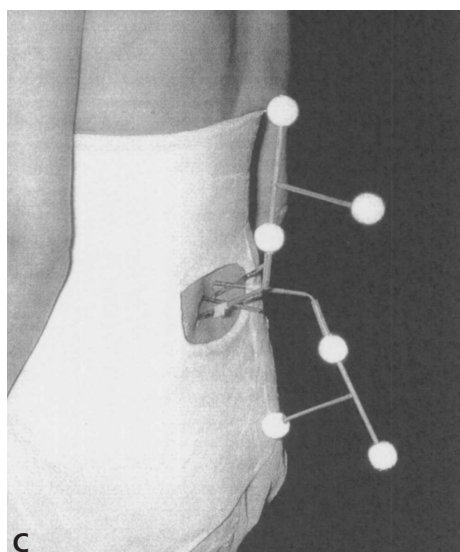


Figure 1. A, The hole created in the plaster casts to permit free mobility to the protruding screw ends. **B,** Rigid attachment of three markers to a proximal screw end and three to a sacral screw end. **C,** The same procedure in a plaster cast with unilateral hip extension.

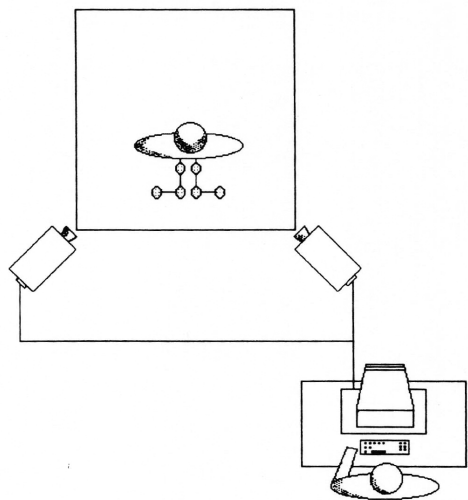
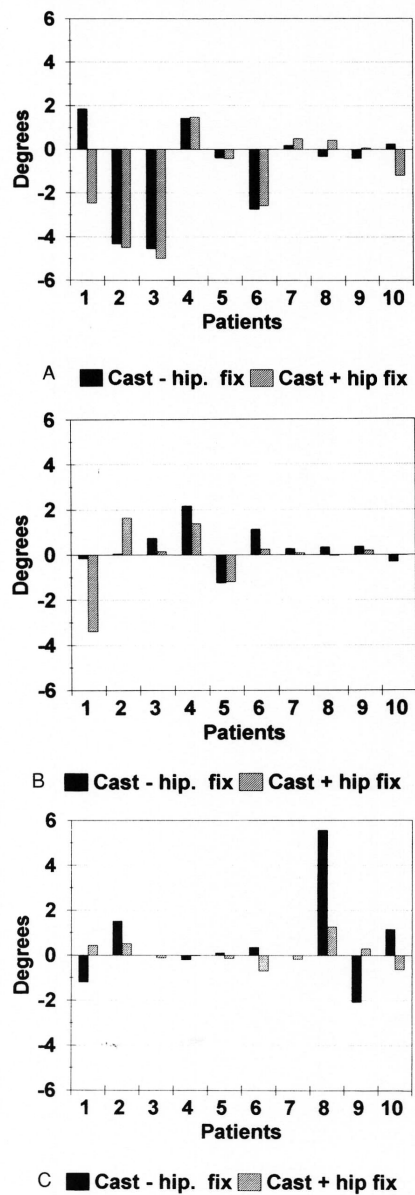


Figure 2. The precision motion analysis system consists of two video cameras connected to a personal computer and creates a measuring volume around the standing patient in which relative movements of the reflective markers during the exercises can be registered.

Figure 3. The differences (Delta) in range of sagittal rotation from both plaster casts to the range obtained in each patient in the condition with no support, during spinal flexion-extension (A), pelvic tilt (B), and walking (C). Note that a negative delta represents an immobilizing effect.



DISCUSSION

The presumed stabilizing effect of a lumbar orthosis on the lumbosacral joint has been the subject of many studies in the past years, and contradictory reports have been published^{1,2,4,7,11}. In a study using Steinmann pins inserted into the spinous process of the fifth lumbar vertebra and the posterior superior iliac spines, Lumsden and Morris¹¹ found that a lumbar orthosis increases rather than decreases lumbosacral rotations during normal levels of walking and concluded that its effect is varied and unpredictable. Buchalter et al⁴ reported that casts and orthoses do not provide sufficient rigidity to adequately immobilize the lumbar spine. With a rigid support, optimal restriction of movement is achieved approximately midway in length and decreases toward the ends⁷. Adding hip fixation to the support thus should provide maximal stabilization of the lumbosacral joint. By means of flexion-extension lateral radiographs, Fidler and Plasmans⁷ found the Baycast spica (Baycast jacket with inclusion of the left thigh) to be consistent in significantly limiting movement at the lumbosacral level in young healthy volunteers. In contrast, recently Axelsson et al² examined the effect of a molded, rigid thoracolumbosacral orthosis with unilateral hip immobilization in patients 1 month after posterolateral lumbosacral fusion.

Using roentgen stereophotogrammetry, they found no significant stabilizing effect by the orthosis on sagittal, vertical, or transverse intervertebral translations, and they suggested that plaster casts restrict gross motions of the trunk rather than exert a stabilizing effect on the intervertebral mobility of the lumbosacral joint^{1,2,10,16,21}.

It has been stated that a three-dimensional technique is required to accurately assess the movements of the intervertebral joints²¹. In the present study, PRIMAS was used with markers applied to transpedicular screws inserted into the fourth or fifth lumbar vertebra and the sacrum of patients with chronic low back pain. The magnitude of angular rotation of the lumbosacral joint in the sagittal, transverse, and frontal planes was assessed.

Considerable rotations were found only in the sagittal plane. This could be explained by the fact that the lumbosacral joint has very little movement concerning axial rotation or lateral bending, but is as mobile as the other levels of the spine in flexion and extension¹⁷.

In the spinal flexion-extension exercise, five patients showed a significant stabilizing effect with one or both plaster casts. In three of these patients, motion had been measured across the L4-S1 segment, where high sagittal rotation was registered in the unrestricted condition (Table 1). In the pelvic tilt exercise, the initial range of motion in eight patients was lower than 1.2 degrees. Compared with the accuracy of the PRIMAS system, this was considered too low to assess further decreases of motion in a reliable way.

In the walking exercise, considerable sagittal rotation was registered, and no significant effect by either of the plaster casts was noted. Although a substantial number of patients showed a decrease in motion in the static measurements, this was not the case under dynamic (walking) test conditions. This interesting observation deserves further attention, although the small number in this patient group and the generally low values of sagittal rotation measured in this study might have limited statistical value.

In the present patient group, small rotations at the lumbosacral level were detected. In another study, larger motion has been reported while measuring with the patient in lateral decubitus position²³. In the present study, measurements in the standing and walking patient were chosen because the immobilizing effect of the plaster cast in common daily life activities was of interest. Using biplanar radiography, several investigators also have reported significantly smaller movements in patients with low back pain and tension signs than those in healthy volunteers, suggesting that the paravertebral muscles in these patients act specifically to splint the lower lumbar levels to reduce or prevent movement and consequent pain^{18,22}. Moreover, the patients in the present study performed the exercises with four protruding 5-mm Schanz screws fixed in their spine, which also could cause reactive splinting of the lumbar spine. This cautious motion in the present patient group, which might limit extrapolation of the data to other patient populations, could also explain why more rotational movement was found during walking than in the other exercises, as it is probably more difficult for the paravertebral muscles to splint the lumbosacral level in a dynamic walking cycle than in a fixed stable position such as maximal flexion-extension or maximal pelvic tilt. This would suggest that dynamic displacements such as those seen in walking provide a more reliable assessment of intervertebral motions and possible instability

than static end displacements, as has been reported earlier^{3,5,20}.

For the PRIMAS system in this experimental set-up, some errors of measurement should be taken into account. First, the fact that the accuracy of the system for static and moving markers proved to be ± 0.3 degrees implies that rotational ranges less than 0.6 degrees should be analyzed with great caution. Secondly, after the placement of a plaster cast, the neutral erect position of the lumbar spine of the patients might have changed. Because total rotational ranges from maximal flexion to extension were analyzed, however, such a change in position probably would not have produced large errors.

Finally, the standardized position, in which the patients were placed in the field of measurement, might have differed somewhat between the exercise sequences. Calculation showed that a deviation of 10 degrees in position produced an error of 1.6% in sagittal rotation, making it an acceptable limitation for the experiments. In this respect, continuous registration of movements during dynamic exercises, e.g., walking, by means of PRIMAS seems to be a promising method for assessing lumbosacral kinematics in patients with chronic low back pain.

Conclusion

In patients with chronic low back pain, plaster casts, with or without unilateral hip immobilization, decrease movement of the lumbosacral joint under static test conditions (spinal flexion-extension). In this study, however, this stabilizing effect did not appear to be present under dynamic conditions (walking).

REFERENCES

1. Axelsson P, Johnsson R, Stromqvist B. Effect of lumbar orthosis on intervertebral mobility. A roentgen stereophotogrammetric analysis. *Spine* 1992;17:678-81.
2. Axelsson P, Johnsson R, Stromqvist B. Lumbar orthosis with unilateral hip immobilization: Effect on intervertebral mobility determined by roentgen stereophotogrammetric analysis. *Spine* 1993;18:876-9.
3. Boden SD, Wiesel SW. Lumbosacral segmental motion in normal individuals: Have we been measuring instability properly? *Spine* 1990;15:571-5.
4. Buchalter D, Kahanovitz N, Viola K, Dorsky S, Nordin M. Three-dimensional spinal motion measurements. Part two: A noninvasive assessment of lumbar brace immobilization of the spine. *J Spinal Disord* 1989;1:284.
5. Dupuis PR, Yong-Hing K, Cassidy JD, Kirkaldy-Willis WH. Radiologic diagnosis of degenerative lumbar spinal instability. *Spine* 1985;10:262-76.
6. Esses SI, Botsford DJ, Kostuik JP. The role of external skeletal fixation in the assessment of low back disorders. *Spine* 1989;14:594-601.
7. Fidler MW, Plasmans CMT. The effect of four types of support on the segmental mobility of the lumbosacral spine. *J Bone Joint Surg [Am]* 1983;65:943-7.
8. Furnee EH. Computer motion analysis systems: The first two decades. Delft University of Technology, 1989 (PhD thesis).
9. Herrmann J. PRIMAS precision motion analysis system: A short description. In: Boeninck U, Nader M, eds. *Gangbildanalyse, stand der Messtechnik und bedeutung fur die Orthopadie-Technik*. Duderstadt: Mecke, Druck, und Verlag, 1990:60-5.
10. Johnsson R, Selvik G, Stromqvist B, Sundin G. Mobility of the lower lumbar spine after posterolateral fusion determined by roentgen stereophotogrammetric analysis. *Spine* 1990;15: 347-50.
11. Lumsden RM, Morris JM. An in vivo study of axial rotation and immobilization at the lumbosacral joint. *J Bone Joint Surg [Am]* 1968;50:1591-1602.
12. Magerl FP. Stabilization of the lower thoracic and lumbar spine with external skeletal fixation. *Clin Orthop* 1984;189: 125-41.
13. Maxwell SE, Delaney HD. *Designing Experiments and Analyzing Data*. Pacific Grove, CA: Brooks/Cole Publishing Company, 1989:568-72.

14. Norton PL, Brown T. The immobilizing efficiency of the back braces: Their effect on the posture and motion of the lumbosacral spine. *J Bone Joint Surg [Am]* 1957;39:111-38.
15. Olerud S, Sjostrom L, Karlstrom G, Hamberg M. Spontaneous effect of increased stability of the lower lumbar spine in cases of severe chronic back pain. The answer of an external transpeduncular fixation test. *Clin Orthop* 1986;203:67-74.
16. Olsson TH, Selvik G, Willner S. Mobility in the lumbosacral spine after fusion studied with the aid of roentgen stereophotogrammetry. *Clin Orthop* 1977;129:181-90.
17. Pearcy MJ, Portek I, Shepherd JE. Three-dimensional xray analysis of normal movement in the lumbar spine. *Spine* 1984;9:294-7.
18. Pearcy MJ, Portek I, Shepherd JE. The effect of low back pain on lumbar spinal movements measured by three-dimensional xray analysis. *Spine* 1985;10:150-3.
19. Perry J. The use of external support in the treatment of low back pain. *J Bone Joint Surg [Am]* 1970;52:1440-2.
20. Pope MH, Frymoyer JW, Krag MH. Diagnosing instability. *Clin Orthop* 1992;279:60-7.
21. Selvik G. Roentgen stereophotogrammetry: A method for the study of the kinematics of the skeletal system. 2nd ed. *Acta Orthop Scand* 1989;60(Suppl):232.
22. Stokes IAF, Wilder DG, Frymoyer JW, Pope MH. Assessment of patients with low back pain by biplanar radiographic measurement of intervertebral motion. *Spine* 1981;6:233-40.
23. Wood KB, Popp CA, Transfeldt EE, Geissele AE. Radiographic Evaluation of instability in spondylolisthesis. *Spine* 1994;19:1697-1703.

3b

The value of a pantaloon cast test in surgical decision making for chronic low back pain patients: a systematic review of the literature supplemented with a prospective cohort study

Willems PC, Elmans L, Anderson PG, Jacobs WCH, van der Schaaf DB, de Kleuver M

ABSTRACT

The results of lumbar fusion in chronic low back pain (LBP) patients vary considerably, and there is a need for proper patient selection. Lumbosacral orthoses have been widely used to predict outcome, however, with little scientific support. The goal of the present study was to determine the value of a pantaloons cast test in selecting chronic LBP patients for lumbar fusion or conservative management. First, a systematic review of the literature was performed in which two independent reviewers identified studies in Medline, Cochrane and Current Contents databases. Three papers met the selection criteria. In the only study with a control group, a significantly better outcome after fusion compared to conservative treatment was found in patients who reported significant pain relief while in a cast (i.e. a positive cast test). The results of lumbar fusion, however, were not significantly different for patients with a positive and those with a negative cast test. In addition to the review, a clinical cohort study of 257 LBP patients who had been allocated to either lumbar fusion or conservative management by a temporary external transpedicular fixation trial, was performed. Prior to allocation, all had undergone a pantaloons cast test. Patients with no history of prior spine surgery and with a positive pantaloons cast test, had a better outcome after lumbar fusion than those treated conservatively ($p=0.002$, χ^2 -test). In patients with previous spine operations the outcomes were poor and the test was of no value. From the literature and the present patient cohort, it was concluded that only in chronic LBP patients without prior spine surgery, a pantaloons cast test with substantial pain relief suggests a favourable outcome of lumbar fusion compared to conservative management. The test has no value in patients who have had previous spine surgery.

INTRODUCTION

Chronic low back pain (LBP) is a frequent clinical problem in western society that merits our concern because of the large number of patients rendered disabled. If conservative measures fail, lumbar fusion may be beneficial, but there is no true consensus in literature regarding its indications and effectiveness^{15,18}. Particularly in patients with degenerative disc disease (DDD) or prior spine surgery, highly variable and unpredictable results have been reported^{10,11,13,14,25,26,34}. As lumbar fusion is a major undertaking, every effort should be made to ensure its successful outcome. Hence, the need for tests that can help select those patients who will benefit from lumbar fusion.

In everyday practice, lumbar orthoses and pantaloons casts are not only used to provide pain relief in low back pain patients^{1,2,7} but also to predict whether lumbar fusion might be indicated. However, this presumed predictive value has not been substantiated by sound scientific support.

The purpose of the current study was to determine the value of a pantaloons cast test in surgical decision-making for chronic LBP patients by performing a systematic review of the literature supplemented with the results of a prospective cohort study.

MATERIALS AND METHODS

Literature search strategy and selection criteria

To obtain all relevant literature, the most common databases of published literature were searched: Medline (1966 to October 2003), Current Contents (1996 to October 2003) and the Cochrane database of randomized controlled trials (2003, issue 1).

No restrictions were made on language or date. Further screening was performed on the titles and abstracts by two independent reviewers. For final selection, articles had to meet the following criteria:

- Studies reported on at least 20 chronic LBP patients, either with or without prior spine surgery, who did not respond to conservative treatment.
- A well-described orthosis or pantaloons cast was applied for stabilization of the lumbosacral region.

- The pain relief of the orthosis or pantaloone cast was assessed preoperatively and compared to the outcome of lumbar fusion with a minimum follow-up of six months.
- Studies on patients with objective neurologic motor deficits or diagnosed as fracture, infectious disease, ankylosing spondylitis, neoplasm, congenital or adolescent idiopathic scoliosis or kyphosis were excluded.

The references from the selected articles were checked in the same way to ensure no relevant articles had been missed. If both reviewers could not reach consensus, a third reviewer was consulted.

Cohort study

Between April 1990 and October 1999, 278 patients with more than six months of incapacitating LBP without objective neurologic motor deficit underwent a temporary external transpedicular fixation (TETF) trial^{12,27,35} as a final diagnostic test. For all patients the indication for lumbar fusion had remained unclear after routine diagnostic tests, including standard radiographs, either MRI or discography, and a pantaloone cast test that was not considered to be a decisive test in surgical decision-making. Thus, the decision to perform lumbar fusion or to continue nonsurgical management was based on the results of the TETF trial. All patients gave informed consent for the TETF procedure and analysis of the patients was approved by an institutional review board.

For the pantaloone cast test, a plaster cast was applied on the standing patient from the nipples to the waist, extending over one leg to the knee. Patients were expected to wear this corset for 6 weeks, with an absolute minimum of three weeks for inclusion in this study. They were encouraged to perform as many daily life activities as possible. Patients were instructed to return earlier for cast removal in case of unbearable pain. At a follow-up visit, generally two weeks after cast removal, the patients were asked whether they had experienced substantial pain relief when wearing the cast. If no significant relief or even a worsening of symptoms had been experienced, a negative test result was noted; in case of significant pain relief, the test was considered positive.

Before the allocation to treatment by the TETF trial, the patients were asked to rate the average pain they had experienced during the previous month on a

horizontal Visual Analogue Scale (VAS) from 0 (no pain) to 100 mm (maximal, unbearable pain). All patients were admitted to hospital for the TETF trial. In case of a significant pain reduction in the fixed position of the TETF compared to the nonfixed position (placebo trial), the result was positive with the advice to perform lumbar fusion³⁵. If pain reduction was absent or not substantial, conservative management was to be continued. At a median follow-up of 76 (15-144) months, patient satisfaction (global self-rating in excellent, good, fair or poor) and pain on the VAS were assessed by a questionnaire that was sent by L.E., who had no previous contact with any of the patients. For analysis, the pain score on VAS at follow-up was subtracted from the initial score before the TETF trial. We defined a decrease of 30% or more from the initial pain to be clinically relevant and scored this as a successful outcome. Patients with less than 30% pain reduction were considered failures. For two of the 278 patients the initial VAS scores were lacking and 19 patients could not be traced for follow-up. Thus, 257 patients remained for analysis (Table 1). Of these 257 patients, 158 had undergone prior spine surgery

Table 1. Patient variables for both treatment groups

	Lumbar fusion (n = 107)	Conservative management (n = 150)
Age (years)	40 ± 8.8	41 ± 8.5
Gender	68 female (64%)	97 female (65%)
Baseline VAS (mean ± sd)	75 ± 16.9	74 ± 15.9
Degenerative Disc Disease	34 (32%)	51 (34%)
Spondylolysis /-olisthesis	3 (3%)	11 (7%)
Prior spine surgery	70 (65%)	88 (59%)
- Postdiscectomy syndrome	29 (27%)	30 (20%)
- Pseudarthrosis	7 (6%)	9 (6%)
- Combination DDD and prior operation	31 (29%)	46 (30%)
- Postchemonucleolysis	2 (2%)	2 (1%)
- Post-trauma	1 (1%)	1 (1%)

while 99 patients had no history of spine surgery. A model was made using a logistic regression to determine which variables contributed to the clinically relevant pain decrease of 30% or more on the VAS score. The scalar variables included were age, duration of symptoms and length of follow-up. The dichotomized variables included were gender, prior spine surgery, pantaloons cast test result and treatment (lumbar fusion or conservative management). Variables with a p-value less than 0.10 were considered to be possible contributors. Two interaction terms were included: prior spine surgery versus treatment and pantaloons cast test result versus treatment. Pearson's Chi Square test was used to calculate statistical significance of differences between groups and correlations were tested with Spearman's Rank Correlation test. A p-value of 0.05 was considered significant.

RESULTS

Literature search

The search yielded 348 references. On the basis of the title, 284 references could be excluded as it was clear that the articles were either a case report, were not related to the lumbar spine, or reported on patients with diagnoses not pertaining to the inclusion criteria. Most of the remaining 64 studies were excluded on the basis of the abstract in which it was evident that lumbar fusion was not performed as treatment, no orthosis or pantaloons cast had been used or had only been used in the postoperative phase. Four articles, all in English, met the inclusion criteria for further review. One study³¹ could not be included because only 15 patients with a preoperative cast test were reported. The three remaining articles and their relevant data are listed in Table 2. Two studies used a pantaloons cast^{22,29}, whereas the other study used rigid orthoses or canvas corsets without hip extension⁵. Markwalder and Reulen²², reported that 23 of 25 (92%) post-discectomy patients had a subjective good outcome after lumbar fusion following a positive plaster cast test. As only test-positive patients undergoing lumbar fusion were included, there were no control groups of negatively tested or conservatively treated patients. Rask and Dall²⁹ included both patients with and those without prior spine surgery. Patients with a positive pantaloons cast test, who subsequently underwent lumbar fusion, did significantly better at about one year follow-up than conservatively treated

Table 2. Results of the selected literature supplemented with the findings of the current clinical study.

	Current study; 2005	Rask and Dall ²⁹ ; 1993	Axelsson et al. ⁵ ; 1995	Markwalder et al. ²² ; 1989
No. of patients	257	45	50	25
Diagnosis	158 prior spine surgery and 99 DDD / spondylolisthesis	17 prior spine surgery and 28 DDD patients	15 prior spine surgery and 35 DDD patients	All prior spine surgery (post-discectomy) patients
Intervention	Pantaloons plaster cast	Fiberglass pantaloons cast	Molded rigid TLSO or canvas corset with plastic posterior reinforcement	Molded plaster body jacket with optional unilateral hip extension
Study design	Both test-positive and test-negative patients included; both lumbar fusion and control group of nonsurgical management	Both test-positive and test-negative patients included; both lumbar fusion and control group of nonsurgical management	Both test-positive and test-negative patients included; all patients fused / no control group of nonsurgical management	Only test-positive patients included; all patients fused / no control group of nonsurgical management
Length of follow-up	Median of 76 (15-144) months	Mean of 12 months with a minimum of 6	All patients 24 months	Mean of 19 months, minimum unknown
Results	Significant better outcome of fusion compared to conservative treatment in patients with a positive cast test and no history of prior spine surgery (p=0.002)	Significant better outcome of fusion compared to conservative treatment in patients with a positive cast test (p=0.016)	No significant correlation of corset test and outcome of lumbar fusion	92% subjective good outcome of fusion in positively tested patients

patients. As only two test-negative patients underwent lumbar fusion, no adequate comparison of outcome between test-positive and test-negative patients could be made. Further, the results for patients with or without prior spine surgery were not stratified. Axelsson et al⁵ included both patients with and those without prior spine surgery, for whom the results were not stratified. No significant difference in outcome after lumbar fusion was found between patients with a positive and those with a negative corset test. As all patients were fused, no control group of nonsurgical treatment was reported.

Cohort study

Of the 257 analyzed patients, 98 experienced significant pain relief when wearing the pantaloons cast, whereas 159 patients had a negative test result. Apart from inconvenience, no complications from wearing the cast were noted. There was no significant correlation between the pantaloons cast test result and the advice from the TETF trial ($\rho=0.01$, $p=0.891$). Based on the TETF, 107 patients actually underwent lumbar fusion: 79 patients underwent an instrumented posterolateral intertransverse process fusion and 28 patients had an anterior interbody fusion. For both operation techniques about 40% of the patients had a clinically relevant improvement. There was a moderate correlation ($\rho=0.62$, $p=0.043$) between overall patient satisfaction and outcome determined by a decrease of pain of 30% or more on the VAS. The distributions of the VAS scores before treatment and at follow-up are shown for all patients in Figures 1 through 4.

According to the logistic regression model, only the interactions between treatment and the pantaloons cast test result ($p=0.079$) and that between treatment and whether the patient had undergone prior spine surgery ($p=0.004$) were factors contributing to successful pain relief in the study population. These interaction relations were tested by a separate Chi Square analysis on clinical outcome x pantaloons cast test result both for patients with and without a history of prior spine surgery. For the group of 99 patients with no prior spine surgery, the patients with a positive pantaloons cast test who underwent a lumbar fusion, had a significantly better outcome than those managed conservatively ($p = 0.002$, Table 3a). Those with a negative test showed no difference in outcome between fusion and conservative treatment. In the group of 158 patients with prior spine surgery no

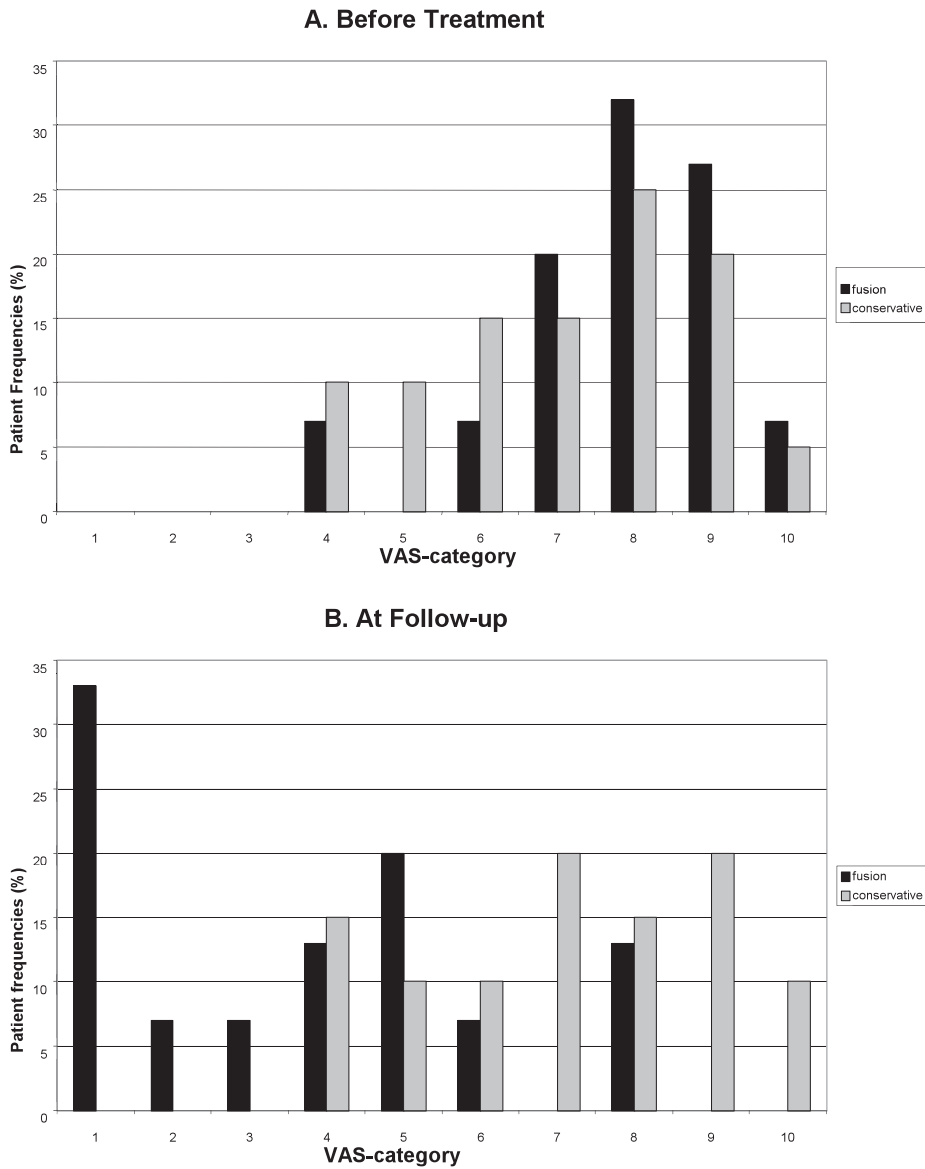


Figure 1. The distribution of the VAS scores before treatment (A) and at follow-up (B) for all patients with no prior spine surgery and a positive plaster cast test (see also groups A and C, Table 3a). The VAS scores have been categorized as follows: 1=0-10; 2=11-20; 3=21-30, etc.

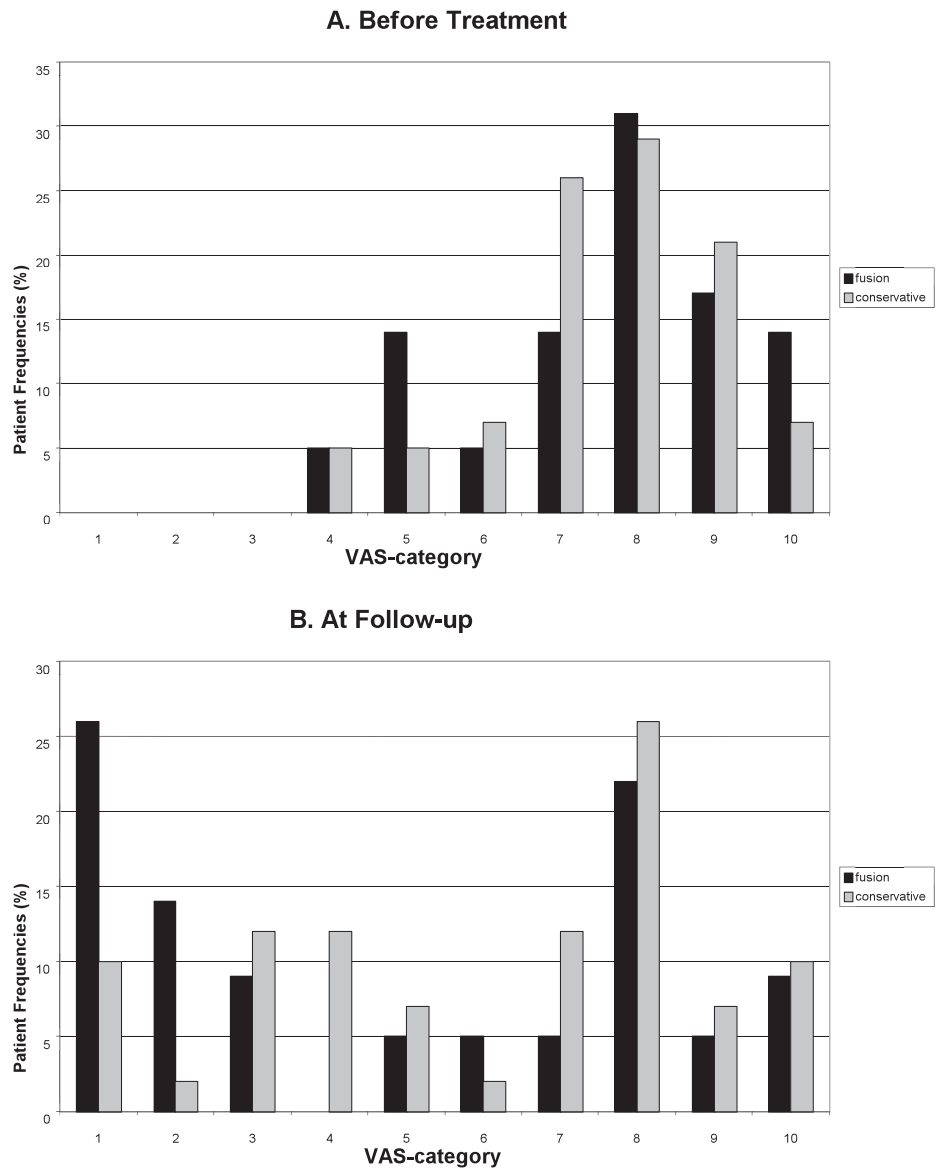


Figure 2. The distribution of the VAS scores before treatment (A) and at follow-up (B) for all patients with no prior spine surgery and a negative plaster cast test (see also groups B and D, Table 3a). The VAS scores have been categorized as follows: 1=0-10; 2=11-20; 3=21-30, etc.

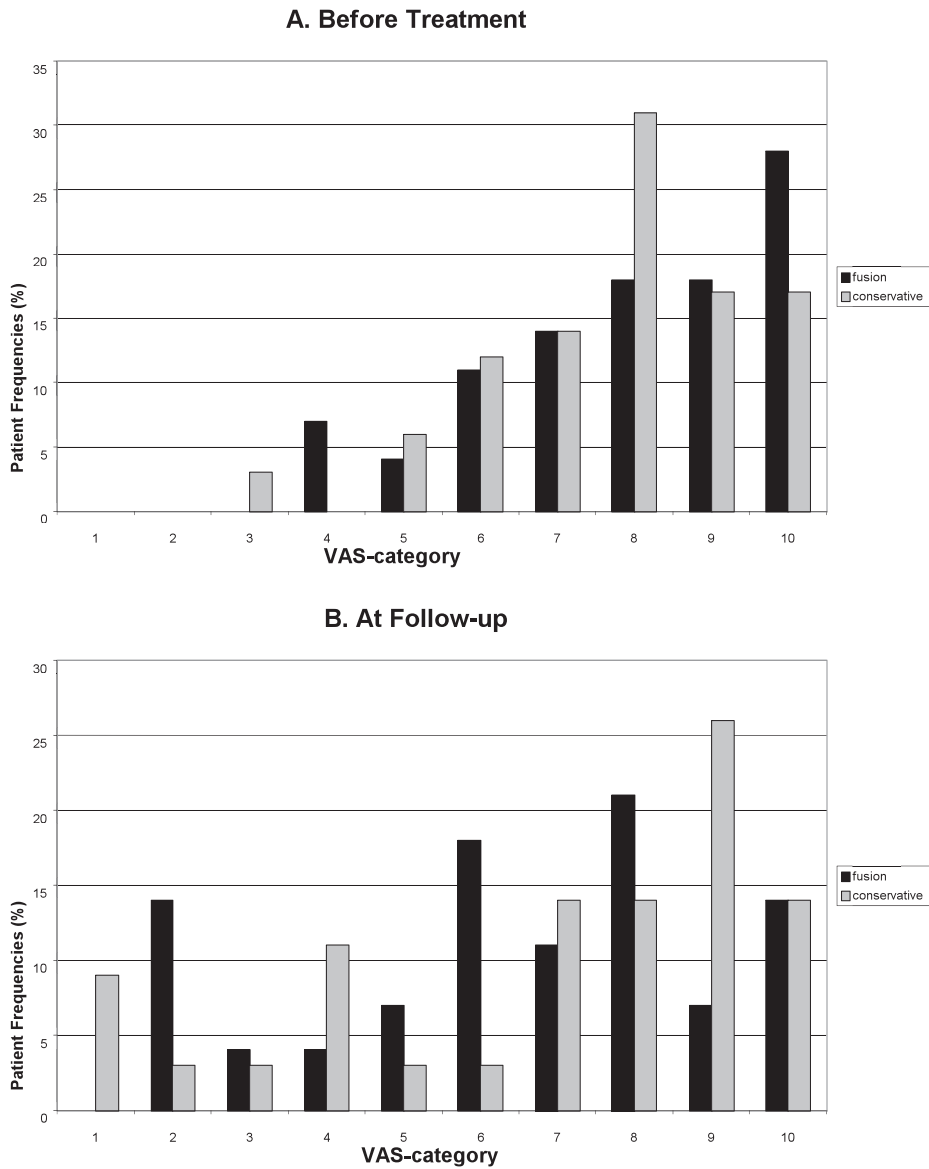


Figure 3. The distribution of the VAS scores before treatment (A) and at follow-up (B) for all patients with prior spine surgery and a positive plaster cast test (see also groups A and C, Table 3b). The VAS scores have been categorized as follows: 1=0-10; 2=11-20; 3=21-30, etc.

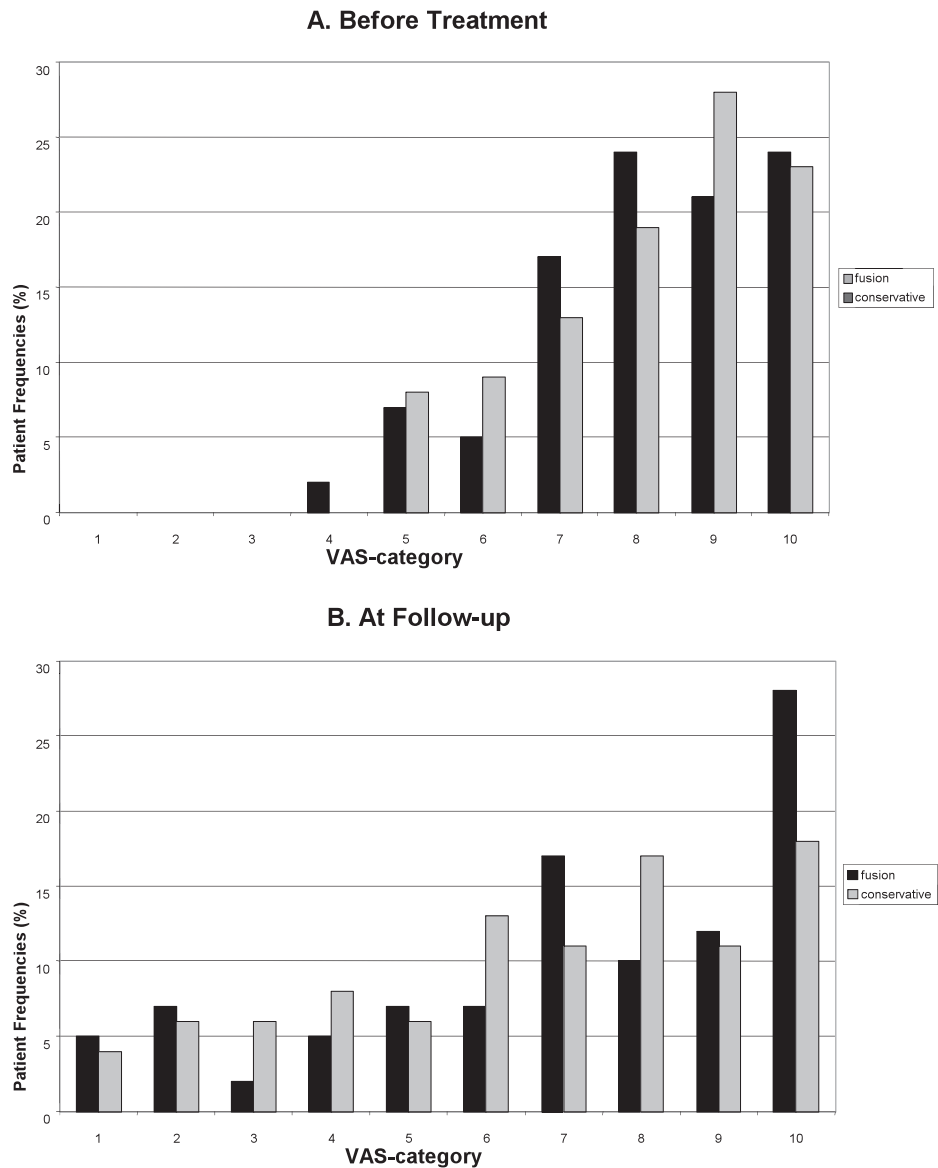


Figure 4. The distribution of the VAS scores before treatment (A) and at follow-up (B) for all patients with no prior spine surgery and a negative plaster cast test (see also groups B and D, Table 3b). The VAS scores have been categorized as follows: 1=0-10; 2=11-20; 3=21-30, etc.

difference in outcome was found (Table 3b). None of the patient groups showed a significant difference in outcome after lumbar fusion between the patients with a positive and those with a negative pantaloon cast test.

Table 3a. Summary of results in the 99 patients without prior spine surgery, classified by treatment and pantaloon cast test result

	Lumbar fusion		Conservative management	
	A: Positive pantaloon cast test	B: Negative pantaloon cast test	C: Positive pantaloon cast test	D: Negative pantaloon cast test
Number of patients	15	22	20	42
Median duration of pain (years + range)	3.0 (0.5 - 10)	3.8 (1.0 – 18)	2.5 (0.5 – 20)	3.3 (0.5 – 22)
Median follow-up (months + range)	66 (40 – 137)	77 (18 – 124)	52 (32 – 115)	69 (30 – 144)
Baseline VAS (mean ± sd)	74 ± 15	72 ± 18	66 ± 17	72 ± 14
Post-treatment VAS (mean ± sd)	32 ± 26	43 ± 34	67 ± 21	56 ± 29
Pain relief of ≥ 30% on VAS#	11 (73%)	13 (59%)	4 (20%)	17 (40%)

#Significant more pain relief of ≥ 30% on VAS in Group A compared to Group C ($p = 0.002$). No significant difference in pain relief between the Groups A and B ($p = 0.373$), B and D ($p = 0.156$) or between the Groups C and D ($p = 0.111$)

Table 3b. Summary of results in the 158 patients with prior spine surgery, classified by treatment and pantaloon cast test result

	Lumbar fusion		Conservative management	
	A: Positive pantaloon cast test	B: Negative pantaloon cast test	C: Positive pantaloon cast test	D: Negative pantaloon cast test
Number of patients	28	42	35	53
Median duration of pain (years + range)	3.5 (1.0 - 20)	4.0 (0.5 – 31)	4.0 (1.0 – 22)	3.0 (1.0 – 21)
Median follow-up (months + range)	53 (18 – 138)	89 (23 – 137)	89 (30 – 136)	89 (23 – 137)
Baseline VAS (mean ± sd)	76 ± 19	77 ± 16	74 ± 17	77 ± 16
Post-treatment VAS (mean ± sd)	61 ± 28	67 ± 29	65 ± 29	63 ± 27
Pain relief of ≥ 30% on VAS#	8 (29%)	12 (29%)	9 (26%)	21 (40%)

#No significant differences in pain relief between any of the groups

DISCUSSION

The use of a plaster cast test in surgical decision-making for chronic LBP patients remains controversial. In search for scientific proof, the present systematic literature review was performed. All three selected papers had different study designs and conclusions. The study by Rask and Dall²⁹, the only report with a conservatively treated control group, concluded that patients with a positive pantaloon cast test who underwent a lumbar fusion had a significantly better outcome than those treated conservatively. Unfortunately, they made no stratification for patients with or without prior spine surgery, whereas the current clinical study shows a considerable difference between these patient groups. In addition, outcome was measured at an

average follow-up of about a year, which, because of the profound general placebo effect after surgery³³, may be rather short, particularly in spine surgery. The study by Markwalder and Reulen²² included postdiscectomy patients with a positive plaster cast test, all of whom underwent lumbar fusion. As there was no control group in the study design, any selective value of the plaster cast could not be determined. In the study by Axelsson et al.⁵, a rigid orthosis or canvas corset without unilateral hip immobilization was used and, as in the current cohort study, no significant difference in outcome for lumbar fusion at the two-year follow-up was registered between patients with a positive and those with a negative corset test. As there was no conservative control group, the study design made it impossible to assess any selective value of a corset test for lumbar fusion versus nonoperative management. Moreover, patients with and without a history of prior spine surgery were not analyzed separately. Finally, the absence of a unilateral hip extension might decrease any stabilizing effect on the lumbosacral joint, and it could be argued that a removable lumbar orthosis or corset is not as effective as a pantaloon plaster cast that is worn continuously.

To supplement the findings of the current literature review, the clinical outcome of 257 chronic LBP patients who either underwent lumbar fusion or conservative management was assessed. In patients with no history of prior spine surgery and a positive pantaloon cast test, the success rate for lumbar fusion was significantly higher than that for conservative management ($p=0.002$) although the number of patients in both subgroups was relatively small (15 and 20 patients, respectively, Table 4a). Within the lumbar fusion treatment group, the outcome between patients with a positive pantaloon cast test was not significantly different from that for patients with a negative test (73% versus 59%, respectively, Table 3a). Thus, a negative pantaloon cast test does not mean that lumbar fusion will not be successful. This implies a low specificity of the test and only a positive test result might be regarded as providing an additional piece of information in favour of lumbar fusion in the individual patient. In patients with prior spine surgery, the success rate for lumbar fusion was low and the pantaloon cast test was of no value. If in these patients equivocal results exist after routine diagnostic tests, great caution with regard to lumbar fusion is warranted. These findings are in accordance with other reports, stating that there are no reliable tests for preoperative selection

of failed back surgery patients and that, consequently, the results of lumbar fusion are discouraging^{16,20,34,36}.

Although we realize that pain is merely one of the many factors related to the problem “chronic low back pain”, the criterion of pain reduction on a Visual Analogue Scale was chosen to assess successful outcome. Despite the inherent inaccuracy in how the patient puts a dot on the line, it is our opinion that, because the patient serves as his own control, this is an objective method to evaluate outcome of treatment in chronic pain²⁸. Because of the heterogeneity in VAS scores within each group, the difference between pain on VAS before allocation and at follow-up was calculated on an individual level and not on a group level. In this way, it was possible to assess clinically relevant improvements in pain. It should be acknowledged that the cut-off level of 30% pain reduction to define success or failure is arbitrary and debatable and one should realize that, as we know from clinical practice, many of the cases defined as success in the present study still suffer from residual low back pain after treatment.

The length of follow-up did not appear to be a contributing factor to pain reduction in the logistic regression ($p=0.86$). Nevertheless, the large range in follow-up and thus the great variety in interval between the pre- and posttreatment pain assessment, should be regarded as a limitation of the present cohort study.

TETF is an invasive test and high incidences of pin tract infections and neurological complications have been reported^{17,21,30}. Therefore, the TETF trial was only used as a final diagnostic test in patients for whom the indication for lumbar fusion remained unclear after routine diagnostic tests. The patient population in the current study might thus be biased because “clear-cut” cases that were indicated for fusion straight away were not included in the trial. This may limit the generalizability of the current study’s findings in other settings and may have led to lower overall results for lumbar fusion. Our aim, however, was not to assess the success of lumbar fusion in chronic low back pain patients, but to evaluate whether a pantaloons cast test can help select those patients who will benefit more from lumbar fusion than from nonsurgical treatment.

Unfortunately, the literature review did not provide a randomized study which would be the best way to evaluate a preoperative selection test. Within the limits of current clinical practice we considered the present cohort study design to be a

reasonable alternative as the pantaloone cast test result had no correlation with the surgical decision-making based on the TETF trial ($\rho=0.01$, $p=0.891$). In this way, whether after a positive or a negative pantaloone cast test, the groups of fused and conservatively managed patients were fairly equal in size and had similar patient characteristics making analysis between these groups more valid.

Founded on the surgeon's personal belief or experience rather than supported by scientific literature, the pain-relieving effect obtained by spinal immobilization in a lumbar orthosis or pantaloone cast is widely used for patient selection in clinical practice. The mechanical basis, however, is poor, as several studies could not determine a stabilizing effect on intervertebral mobility^{3,4,37}. In fact, radiographic studies have even suggested that motions at lower lumbar levels actually increase when orthoses are worn^{4,32}. The main effect produced by a corset is a reduction in gross motion of the trunk^{3,6,19} rather than the fixation at one individual level. This non-specific effect makes comparison to a single level fixation questionable which is supported by the fact that there appeared to be no correlation between the results of the TETF trial and the pantaloone cast test. The pain relief provided by a corset may perhaps not be caused by motion restriction but by an increase in passive trunk stiffness of up to 40%^{8,9,23} with a possible general unloading effect on the trunk structures and reduced intradiscal pressures²⁴. Such pain relief caused by this unloading effect does not appear to be specific, as we found that a negative pantaloone cast test does not at all exclude the possibility that the patient will benefit from lumbar fusion.

CONCLUSIONS

Chronic LBP patients with no history of prior spine surgery who experience substantial pain relief in a pantaloone cast have a significantly better outcome after lumbar fusion compared to conservative management. A negative test result, however, has no value in surgical decision-making. The test is of no use in patients with a history of prior spine surgery and great caution regarding lumbar fusion is warranted in these patients as the results are very poor. Apart from inconvenience, the pantaloone cast test has a negligible complication rate. Despite its worldwide use, to date, only one other controlled study to determine the value of the test has been published.

REFERENCES

1. Ahlgren SA Hansen T (1978) The use of lumbosacral corsets prescribed for low back pain. *Prosthet Orthot Int* 2, 2: 101-104.
2. Alaranta H Hurri H (1988) Compliance and subjective relief by corset treatment in chronic low back pain. *Scand J Rehabil Med* 20, 3: 133-136.
3. Axelsson P, Johnsson R, Stromqvist B (1992) Effect of lumbar orthosis on intervertebral mobility. A roentgen stereophotogrammetric analysis. *Spine* 17, 6: 678-681.
4. Axelsson P, Johnsson R, Stromqvist B (1993) Lumbar orthosis with unilateral hip immobilization. Effect on intervertebral mobility determined by roentgen stereophotogrammetric analysis. *Spine* 18, 7: 876-879.
5. Axelsson P, Johnsson R, Stromqvist B, Nilsson LT, Akesson M (1995) Orthosis as prognostic instrument in lumbar fusion: no predictive value in 50 cases followed prospectively. *J Spinal Disord* 8, 4: 284-288.
6. Buchalter D, Kahanovitz N, Viola K, Dorsky S, Nordin M (1988) Three-dimensional spinal motion measurements. Part 2: A noninvasive assessment of lumbar brace immobilization of the spine. *J Spinal Disord* 1, 4: 284-286.
7. Calmels P Fayolle-Minon I (1996) An update on orthotic devices for the lumbar spine based on a review of the literature. *Rev Rhum Engl Ed* 63, 4: 285-291.
8. Cholewicki J, Alvi K, Silfies SP, Bartolomei J (2003) Comparison of motion restriction and trunk stiffness provided by three thoracolumbosacral orthoses (TLSOs). *J Spinal Disord Tech* 16, 5: 461-468.
9. Cholewicki J, Juluru K, Radebold A, Panjabi MM, McGill SM (1999) Lumbar spine stability can be augmented with an abdominal belt and/or increased intra-abdominal pressure. *Eur Spine J* 8, 5: 388-395.
10. Deyo RA, Cherkin D, Conrad D, Volinn E (1991) Cost, controversy, crisis: low back pain and the health of the public. *Annu Rev Public Health* 12, 141-156.
11. Dunsker SB (1990) Lumbar spine stabilization: indications. *Clin Neurosurg* 36, 147-158.
12. Esses SI, Botsford DJ, Kostuik JP (1989) The role of external spinal skeletal fixation in the assessment of low-back disorders. *Spine* 14, 6: 594-601.
13. Esses SI Moro JK (1993) The value of facet joint blocks in patient selection for lumbar fusion. *Spine* 18, 2: 185-190.

14. Fritzell P, Hagg O, Wessberg P, Nordwall A (2001) 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 26, 23: 2521-2532.
15. Gibson JN, Grant IC, Waddell G (1999) The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 24, 17: 1820-1832.
16. Jackson RK, Boston DA, Edge AJ (1985) Lateral mass fusion. A prospective study of a consecutive series with long-term follow-up. *Spine* 10, 9: 828-832.
17. Jeanneret B, Jovanovic M, Magerl F (1994) Percutaneous diagnostic stabilization for low back pain. Correlation with results after fusion operations. *Clin Orthop* 304: 130-138.
18. Krismer M (2002) Fusion of the lumbar spine. A consideration of the indications. *J Bone Joint Surg Br* 84, 6: 783-794.
19. Lantz SA, Schultz AB (1986) Lumbar spine orthosis wearing. I. Restriction of gross body motions. *Spine* 11, 8: 834-837.
20. Lehmann TR, Spratt KF, Tozzi JE et al (1987) Long-term follow-up of lower lumbar fusion patients. *Spine* 12, 2: 97-104.
21. Lund T, Oxland TR, Nydegger T, Schlenzka D, Laine T, Heini P (2002) Is there a connection between the clinical response after an external fixation test or a subsequent lumbar fusion and the pre-test intervertebral kinematics? *Spine* 27, 23: 2726-2733.
22. Markwalder TM, Reulen HJ (1989) Diagnostic approach in instability and irritative state of a "lumbar motion segment" following disc surgery--failed back surgery syndrome. *Acta Neurochir (Wien)* 99, 1-2: 51-57.
23. McGill S, Seguin J, Bennett G (1994) Passive stiffness of the lumbar torso in flexion, extension, lateral bending, and axial rotation. Effect of belt wearing and breath holding. *Spine* 19, 6: 696-704.
24. Nachemson A, Morris JM (1964) In vivo measurements of intradiscal pressure. Discometry, a method for the determination of pressure in the lower lumbar discs. *J Bone Joint Surg Am* 46, 1077-1092.
25. Nachemson AL (1985) Fusion for low back pain and sciatica. *Acta Orthop Scand* 56, 4: 285-286.
26. O'Beirne J, O'Neill D, Gallagher J, Williams DH (1992) Spinal fusion for back pain: a clinical and radiological review. *J Spinal Disord* 5, 1: 32-38.

27. Olerud S, Karlstrom G, Sjostrom L (1988) Transpedicular fixation of thoracolumbar vertebral fractures. *Clin Orthop* 227, 44-51.
28. Price DD, McGrath PA, Rafii A, Buckingham B (1983) The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain* 17, 1: 45-56.
29. Rask B, Dall BE (1993) Use of the pantaloone cast for the selection of fusion candidates in the treatment of chronic low back pain. *Clin Orthop* 288: 148-157.
30. Soini JR, Seitsalo SK (1993) The external fixation test of the lumbar spine. 30 complications in 25 of 100 consecutive patients. *Acta Orthop Scand* 64, 2: 147-149.
31. Tokuhashi Y, Matsuzaki H, Sano S (1993) Evaluation of clinical lumbar instability using the treadmill. *Spine* 18, 15: 2321-2324.
32. Tuong NH, Dansereau J, Maurais G, Herrera R (1998) Three-dimensional evaluation of lumbar orthosis effects on spinal behavior. *J Rehabil Res Dev* 35, 1: 34-42.
33. Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE (1994) The importance of placebo effects in pain treatment and research. *JAMA* 271, 20: 1609-1614.
34. Turner JA, Ersek M, Herron L et al (1992) Patient outcomes after lumbar spinal fusions. *JAMA* 268, 7: 907-911.
35. van der Schaaf DB, van Limbeek J, Pavlov PW (1999) Temporary external transpedicular fixation of the lumbosacral spine. *Spine* 24, 5: 481-484.
36. Wetzel FT, LaRocca SH, Lowery GL, Aprill CN (1994) The treatment of lumbar spinal pain syndromes diagnosed by discography. Lumbar arthrodesis. *Spine* 19, 7: 792-800.
37. Willems PC, Nienhuis B, Sietsma M, van der Schaaf DB, Pavlov PW (1997) The effect of a plaster cast on lumbosacral joint motion. An in vivo assessment with precision motion analysis system. *Spine* 22, 11: 1229-1234.

4a

Provocative discography and lumbar fusion Is preoperative assessment of adjacent discs useful?

Willems PC, Elmans L, Anderson PG, van der Schaaf DB, de Kleuver M

Spine 2007;32:1094-1099

ABSTRACT

Study Design. A cohort study of clinical outcomes of lumbar fusion patients with preoperative assessment of adjacent levels by provocative discography.

Objective. To evaluate whether the preoperative status of the adjacent discs, as determined by provocative discography, has an impact on the clinical outcome of lumbar fusion in chronic low back pain (LBP) patients.

Summary of Background Data. The results of lumbar fusion in chronic LBP patients vary considerably and are hard to predict. It is believed that degenerative levels adjacent to a fused spinal segment may be a cause of continuing pain. In this respect it is important to know whether preoperative degenerative or symptomatic adjacent levels have an adverse effect on patient outcomes after lumbar fusion.

Methods. In 197 patients with an equivocal indication for lumbar fusion (two-thirds were patients with prior spine surgery), the decision for either lumbar fusion or conservative management was determined by a temporary external transpedicular fixation (TETF) trial. During the diagnostic work-up, all patients had undergone provocative discography that included the assessment of the discs adjacent to the intended fusion levels. The individual changes in pain on a visual analogue scale, assessed before treatment and at follow-up, and patient satisfaction were the measures of outcome.

Results. In the 82 patients who underwent a lumbar fusion, no difference in outcome was found between those patients with degenerative or symptomatic discs adjacent to the fusion and those with normal adjacent discs.

Conclusion. In this cohort study of chronic LBP patients with an uncertain indication for lumbar fusion, the preoperative status of adjacent levels as assessed by provocative discography did not appear to be related to the clinical outcome after fusion.

INTRODUCTION

“Diagnostic disk puncture with injection of opaque medium demonstrates disk ruptures and protrusions and tells if the patient’s symptoms originate from the punctured disk. The method seems to be of great practical value.” (Lindblom, 1948)¹.

Chronic low back pain (LBP) is a complex clinical and socio-economic problem in modern society leading to a high use of health services². As a routine work-up often fails to provide a clear diagnosis, the evaluation of LBP patients can be both challenging and frustrating. Lumbar fusion may be beneficial for some patients, but its effectiveness has been reported to be highly variable and hard to predict³⁻⁵. Furthermore, when a rigid motion segment is created, extra stresses are exerted on the adjacent motion segments with the potential risk of accelerated degeneration of adjacent levels and increasing pain⁶⁻⁸. Therefore, patients with a fusion that ends adjacent to an already degenerative or painful disc may be expected to have more pain and thus poorer clinical results at follow-up than those patients whose adjacent discs are normal⁹. For this reason, when lumbar fusion is proposed for chronic LBP patients, it is considered wise either to include symptomatic adjacent discs in the fusion or to refrain from any fusion at all¹⁰.

Imaging modalities, such as plain radiographs, magnetic resonance imaging (MRI) or computerized tomography (CT), are quite sensitive in detecting disc degeneration but cannot confirm whether a disc is symptomatic and relevant in the patient’s pain syndrome¹¹⁻¹³.

Provocative lumbar discography is a physiologic test with which it can be determined whether a disc is painful upon injection. Even if degenerative signs are absent on imaging, discography can reveal the presence of a symptomatic rim lesion^{14,15}. A severe, acute exacerbation of the patient’s usual pain provoked by discography is assumed by some spine specialists to be diagnostic of a clinically relevant pain generator at the disc itself, provided that the injection of a control disc does not reproduce familiar pain¹⁶⁻²⁰. When planning the extent of a lumbar fusion, provocative discography can be used to verify whether the discs adjacent to the intended fusion level(s) do not provoke pain and could be expected to withstand the added stress after fusion.

The goal of the present study was to determine whether the preoperative status of adjacent discs, as assessed by provocative discography, has any impact on the clinical outcome of a cohort of chronic LBP patients in whom the decision to perform lumbar fusion was not based on the discography test itself.

MAREIALS AND METHODS

From April 1990 to October 1999, 209 patients for whom the indication for a lumbar fusion remained uncertain after routine diagnostic tests underwent a temporary external transpedicular fixation (TETF) trial²¹⁻²³ as the final, decisive test. Patients with an unequivocal indication for lumbar fusion after routine diagnostic testing were excluded from the study. All patients had suffered from incapacitating low back pain for more than one year and had no objective neurologic motor deficit. Routine diagnostic tests comprized standard radiographs and provocative discography that included the assessment of discs adjacent to the level(s) intended for fusion. The eventual decision whether to perform lumbar fusion or to continue nonsurgical management was determined by the TETF trial for which all patients gave informed consent. The analysis of the patients was approved by an institutional review board.

Lumbar discography was performed in the Radiology Department on an outpatient basis by orthopaedic residents affiliated with the spine unit. No standard sedatives or prophylactic antibiotics were used. With the patient in the left lateral decubitus position, the technique described by McCulloch and Waddell²⁴ was used. The skin was prepared with Betadine and draped in the routine manner. After infiltration of the skin with local anaesthetic (Lidocain 1%), a stiletted double needle²⁵ (18-gauge outer and 22-gauge inner solid needle) was advanced into the intended disc space. Correct placement of the needle tip into the centre of the disc was verified by biplanar fluoroscopy. In case the disc centre could not be reached properly, an open, more flexible 22-gauge needle was advanced through the outer needle to reach the centre. Subsequently, Isovist®, a water-soluble nonionic contrast agent (Schering AG, Berlin, Germany) was injected. Next to the disc(s) that was suspected to be the source of pain in the patient based on plain radiographs or MRI, generally two adjacent discs (1 proximal and 1 distal level or 2 proximal levels in case L5-S1 was

suspect) were tested. Attention was paid to whether the patient experienced pain during injection. The patient was asked about the extent and location of the pain and whether this pain was similar to his/her usual pain (i.e. concordant pain). The injection endpoint was defined by manual resistance. The morphological pattern of the disc was described according to the criteria of Adams²⁶ in which the discogram types I and II were scored as normal, whereas the types III-V were considered to be degenerative²⁰. For each disc separate needles were used. One hour after the procedure, patients were sent home with the instruction to return immediately in case of aggravated back pain or fever within the next few weeks²⁷. Routine follow-up was at 6 to 12 weeks after the procedure.

All patients were admitted to hospital for the TETF trial and were asked to rate their pain on a horizontal visual analogue scale (VAS) ranging from 0 (no pain) to 100 mm (maximal, unbearable pain). The levels considered for fusion were immobilized by TETF. In case of a significant pain reduction in the fixed position compared to the nonfixed state (=placebo-trial)²³, the result of the TETF trial was considered to be positive and lumbar fusion of the intended level(s) was performed. If pain reduction was absent or not substantial, conservative management with commonly practiced measures, such as physical therapy, manipulations, rehabilitation programs, analgesics, lumbar orthoses or nerve root blocks was to be continued. A standardised nonoperative treatment was not advocated because we feel that commonly used conservative methods tailored to the individual patient portray the real life situation for the average chronic LBP patient²⁸. At follow-up, a questionnaire was sent to the patients by L.E. who had no previous contact with any of the patients. Pain on the VAS and patient satisfaction (global self-rating of outcome in excellent, good, fair or poor) were thus assessed in both the fused and conservatively treated patients. In every patient the pain score at follow-up was subtracted from the initial pain score. We considered an individual decrease in pain of 30% or more of the initial pain to be a clinically relevant improvement and defined this as a successful outcome. Patients with less than 30% pain reduction were considered failures. For analysis, the discographic findings of all patients with at least one injected disc adjacent to the levels that had been immobilized by TETF were reviewed and scored as follows: If provocation of these adjacent disc(s) had produced concordant pain, the discography result was considered negative (which would mean that fusion is not

recommendable). If injection into these disc(s) had provoked unfamiliar pain or no pain at all, the discography result was positive. From one of the original 209 patients the initial data were lost and 11 cases could not be traced for follow-up: Two patients with symptomatic adjacent discs and eight patients with asymptomatic discs who had received conservative treatment and the only fused lost-to-follow-up case was a patient with asymptomatic discs at discography.

Thus, 197 patients (63 male, 134 female) with a mean age of 40 ± 8.5 years remained for analysis at a mean follow-up of 80 ± 35 months (range 15-144 months; Table 1). Based on the TETF results, 115 patients were managed

Table 1. Diagnosis of all analyzed patient groups (number of patients)

	Fusion and symptomatic adjacent discs (n=22)	Fusion and asymptomatic adjacent discs (n=60)	Conservative management (n=115)
Degenerative Disc Disease	4	20	39
Spondylolysis /-olisthesis	2	3	7
Prior Spine Surgery	16	37	69
- Postdiscectomy	5	18	16
- Pseudarthrosis	1	6	7
- Combination DDD and prior operation	7	12	43
- post-chemonucleolysis	3	0	2
- Post-trauma	0	1	1

conservatively and 82 patients underwent a lumbar fusion: 24 patients had an anterior interbody fusion (stand-alone or circumferential) and 58 patients received an instrumented posterolateral intertransverse process fusion. For statistical analysis, the variables to be tested (either dichotomized or ordinally scaled) were age, gender, previous spine surgery, length of follow-up, preoperative VAS pain score, provocative discography result of adjacent discs, treatment (lumbar fusion or conservative), and clinical outcome. A logistic regression model was made to determine which variables contributed significantly to a clinically successful

outcome. Pearson's Chi-Square test was used to calculate statistical significance of differences between groups. For insight into the magnitude of effect, odds ratios (OR) with confidence intervals (CI) were calculated for any parameter with a p-value of 0.05 or less. Correlations were tested with Spearman's Rank Correlation test. A p-value of 0.05 was considered to be significant.

RESULTS

No serious complications related to the discography procedure were registered. If degenerative discs are considered to be painful and normal discs to be asymptomatic, morphology predicted the discographic pain response in about 75% of the patients (21 painful and degenerative discs, 126 asymptomatic discs with normal morphology, Table 2). About two-thirds of the degenerative adjacent discs were painful as compared to a quarter of the adjacent discs with normal morphology ($p = 0.000$ chi-squared test, OR 6.2, CI_{95} 2.7-13.9; Table 2). There was

Table 2. Cross-tabulation of the morphologic (Adams' criteria) and the symptomatic status of the adjacent discs assessed by provocative discography in all 197 patients

	Symptomatic	Not symptomatic	Total number of patients
Degenerative	21	11	32
Normal morphology	39	126	165
Total number of patients	60	137	197

no correlation ($\rho=0.1$, $p=0.332$) between the result of provocative discography and the TETF trial result. According to the logistic regression model, only treatment (lumbar fusion versus conservative management) appeared to be a contributing factor to successful pain relief ($p=0.012$, OR 2.2, CI_{95} 1.18-3.9). One can conclude that a patient in this study who underwent a fusion operation had approximately twice as much chance of having at least a 30% reduction in pain as a patient who was treated conservatively. The other variables in the logistic regression did not contribute significantly to having at least a 30% reduction in pain: age ($p=0.48$),

gender ($p=0.19$), preoperative VAS score ($p=0.83$), previous spine surgery ($p=0.06$), interaction previous spine surgery and treatment ($p=0.98$), provocative discography ($p=0.996$) and length of follow-up ($p=0.48$).

For the 82 fused patients the distributions of the VAS scores before lumbar fusion and at follow-up are shown in Figure 1. The success rate of those fused with asymptomatic adjacent discs was equal to the success rate of those fused with painful adjacent discs (45% versus 45%, respectively, $p=0.583$, Table 3). When considering

Table 3. Variables and clinical outcome of the 82 patients who underwent lumbar fusion classified by the result of provocative discography of the adjacent segments

	Asymptomatic adjacent disc(s)	Symptomatic adjacent disc(s)
Number of patients	60	22
Mean age (years)	39 ± 7.8	39 ± 8.5
Gender (Male / Female)	18 M / 42 F	6 M / 16 F
Mean follow-up in months (range)	80 ± 36 (15-137)	76 ± 38 (18-144)
Mean initial VAS	76 ± 16	72 ± 18
Mean VAS at follow-up	53 ± 33	52 ± 29
Patients with pain relief of $\geq 30\%$ on VAS	27 (45%)	10 (45%)

the effect of discographic morphology on the clinical outcome for lumbar fusion between those patients with degenerative adjacent discs versus those with normal adjacent discs, the Pearson's Chi Square test was not significant ($p=0.072$). Stratification for patients with or without prior spine surgery did not reveal any effect of discography result on outcome (Pearson's Chi Square: no prior surgery, $p=0.847$; prior surgery, $p=0.705$). The method of fusion, whether anterior interbody fusion or instrumented posterolateral fusion, had no effect on outcome (success rates of 46% versus 45%, respectively). The outcome by percentage of VAS improvement and the 4 categories of patient satisfaction (excellent, good, fair or poor) appeared to correlate well ($\rho=0.73$, $p=0.000$). Patient satisfaction, which we consider to be a sufficient quality indicator rather than another validated measure of health outcome, also did not delineate differences between the two operative groups ($p=0.266$).

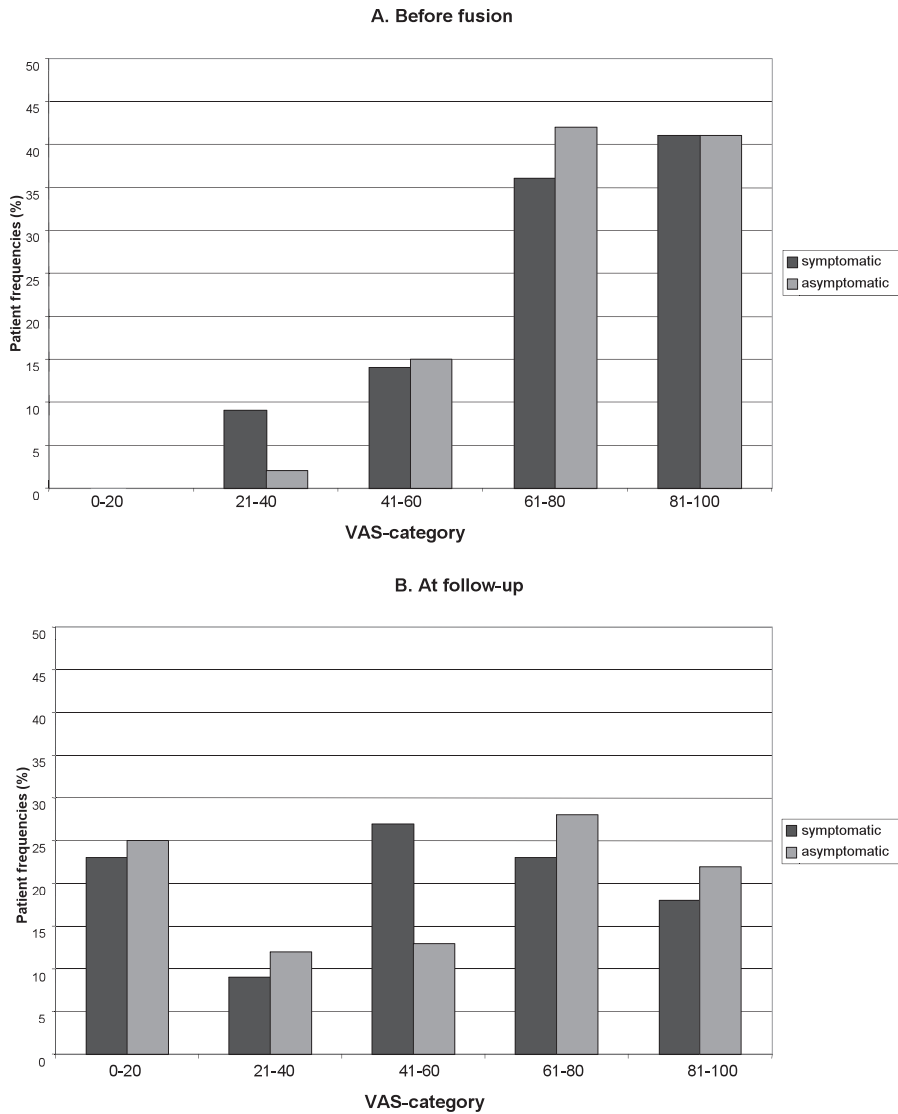


Figure 1. The distribution of the VAS scores before lumbar fusion (A) and at follow-up (B) for all fused patients with symptomatic and asymptomatic adjacent discs at preoperative discography, respectively (see also Table 3). The VAS scores have been categorized as follows: 0-20; 21-40; 41-60; 61-80 and 81-100.

DISCUSSION

There is no true consensus in the management of adjacent segments in spinal fusion. Some authors advocate to limit the number of segments fused as much as possible because two- or three-level fusions have a proven higher rate of pseudarthrosis with less function^{7,29,30} and lower patient satisfaction compared to single level fusions³. Others suggest that surgical fusion of lumbar vertebrae increases the load on adjacent discs which may accelerate degeneration at these transitional levels^{6,31-34}. However, in a long-term follow-up study following surgery for lumbar degenerative disc disease, no correlation between radiographic signs of degeneration at transitional levels and clinical outcome or functional impairment was reported³⁴. In a recent retrospective study Throckmorton et al. reviewed the outcome after lumbar fusion for degenerative instability in 25 patients³⁵. They reported no adverse impact on clinical outcome when the lumbar fusion ended adjacent to a degenerative motion segment as determined from preoperative morphologic MRI data. As a criticism one might argue that radiographically degenerative lumbar discs or so-called black discs on MRI do not appear to have clinical relevance as they are observed in up to a third of asymptomatic individuals between 20 and 39 years of age^{12,36-38}.

Because degeneration itself appears to be related to the process of aging, whereas symptomatic degeneration may point to deterioration³⁶, one could assume that physiologic data from provocative discography would reflect the clinical scenario better in chronic LBP patients. This more reliable preoperative analysis of planned future transitional levels should lead to a better prediction of the clinical outcome of lumbar fusion³⁹. In the present study, however, clinical outcome did not appear to be affected by the preoperative status of adjacent levels as assessed by provocative discography. There are, however, methodological limitations of the present study that should be addressed.

Provocative discography is a controversial test and its validity has not been established consistently in literature⁴⁰. It has been shown that healthy, previously pain-free patients can also experience pain from a discogram and the discographic pain response appears to be highly variable in patients with chronic pain and somatization disorders⁴¹. It should be stressed that discography is not a test at the

level of a gold standard to identify whether a disc is truly a clinically significant pain generator in a chronic LBP patient. It is no more than an assumption frequently given to the discography result which can be compared to clinical outcome.

In the present study, provocative discography was executed by residents who had performed a sufficient number of discography procedures under supervision to be judged capable of executing the procedure alone. Nevertheless, it has been documented that false test results increase with poor injection technique implying that the performance and evaluation of the test by residents may be less reliable than by experienced discographers who perform discography on a routine base.

The decision of which levels were to be fixed by TETF was based on the combination of conventional radiographs, MRI/CT and provocative discography that included the assessment of generally two adjacent discs. In this way, in the individual patient, an originally adjacent disc that was symptomatic on injection, may have become a suspect disc and subsequently been included in the TETF (=proposed fusion) levels. This may have resulted a bias in the selection of both fusion groups. For the present protocol, adjacent levels were defined retrospectively and the discographic responses of only these levels, truly adjacent to TETF (=fusion), were studied. Consequently, in some patients only one adjacent level could be scored for analysis, which may limit the interpretation of the present study's findings.

However, in contrast to most studies, the design of the present study had the advantage that the decision to perform lumbar fusion was not based on the results of adjacent level discography itself (as defined retrospectively) but on the TETF results, which were not correlated ($\rho=0.1$, $p=0.332$). In this way, a group of fused patients with symptomatic or degenerative adjacent discs and a group of fused patients with normal adjacent discs could be compared.

Because TETF is an invasive diagnostic procedure with a considerable risk of pin tract infection and neurological complications⁴²⁻⁴⁴, only patients for whom routine diagnostic testing was equivocal were included. Such a selection manner is, of course, prone to bias by the judgment of the individual clinician. Moreover, whether the TETF trial was positive (sufficient pain relief to decide for lumbar fusion), was determined by the physician who coordinated the TETF trial (D.S.). For his decisions there is no validity other than clinical experience. These methods of selection limit the generalizability of the present study to settings in which patients

are selected for lumbar fusion in a different manner. The aim, however, was not to assess the success of fusion in chronic LBP patients but to evaluate whether the clinical outcome of lumbar fusion is affected by the preoperative status of adjacent discs as determined by provocative discography.

Although pain is only one aspect of the problem of “chronic low back pain”, it remains the principal indication for operative treatment. For this reason, the criterion of significant pain reduction on a VAS was chosen to define successful outcome. The graphic representation of pain over time by the VAS has been shown to be well suited for longitudinal outcomes assessment⁴⁵⁻⁴⁸. We chose for the change in pain per individual to define outcome because of the heterogeneity in VAS scores within each patient group. It should be acknowledged that the choice for the cut-off level of 30% improvement in pain to define success or failure is arbitrary and debatable as it implies (as we also know from current clinical practice) that many of the patients defined as a success in the present study still suffer from residual LBP after treatment.

The spread in time between the pre- and posttreatment pain assessments could be a confounding factor (in both fusion groups there were patients with less than 24 months of follow-up). In the logistic regression, however, length of follow-up was not found to contribute to the clinical outcome of successful pain relief ($p=0.48$).

The total number of 82 fused patients in the present study is small and could lead to a potential type II error. A sample size calculation (power=0.8, alpha set at 0.05) showed that 67 patients are sufficient to show a 15% difference in the rate of success between the two groups (a difference less than 15% would be of little clinical relevance). Therefore, we feel confident that the results presented here justify our critical view concerning the value of evaluation of adjacent levels by provocative discography prior to spinal fusion.

CONCLUSION

In the present cohort of chronic LBP patients the preoperative status of adjacent discs, as determined by provocative discography, did not have any impact on the clinical outcome after lumbar fusion.

REFERENCES

1. Lindblom K. Diagnostic puncture of intervertebral discs in sciatica. *Acta Orthop.Scand.* 1948;17:231-9.
2. Hart LG, Deyo RA, Cherkin DC. Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine* 1995;20:11-9.
3. Turner JA, Ersek M, Herron L et al. Patient outcomes after lumbar spinal fusions. *JAMA* 1992;268:907-11.
4. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 1999;24:1820-32.
5. Krismer M. Fusion of the lumbar spine. A consideration of the indications. *J.Bone Joint Surg.Br.* 2002;84:783-94.
6. Chow DH, Luk KD, Evans JH et al. Effects of short anterior lumbar interbody fusion on biomechanics of neighboring unfused segments. *Spine* 1996;21:549-55.
7. Herkowitz HN, Abraham DJ, Albert TJ. Management of degenerative disc disease above an L5-S1 segment requiring arthrodesis. *Spine* 1999;24:1268-70.
8. Lee CK. Accelerated degeneration of the segment adjacent to a lumbar fusion. *Spine* 1988;13:375-7.
9. Buttermann GR, Heithoff KB, Ogilvie JW et al. Vertebral body MRI related to lumbar fusion results. *Eur.Spine J.* 1997;6:115-20.
10. Cauchoix J, David T. [Lumbar arthrodesis: results after more than 10 years]. *Rev.Chir Orthop.Reparatrice Appar.Mot.* 1985;71:263-8.
11. Brodsky AE, Binder WF. Lumbar discography. Its value in diagnosis and treatment of lumbar disc lesions. *Spine* 1979;4:110-20.
12. Boden SD, Davis DO, Dina TS et al. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J.Bone Joint Surg.Am.* 1990;72:403-8.
13. Wiesel SW, Tsourmas N, Feffer HL et al. A study of computer-assisted tomography. I. The incidence of positive CAT scans in an asymptomatic group of patients. *Spine* 1984;9:549-51.
14. Mooney V. Lumbar discography. *Spine* 1996;21:1479.
15. Osti OL, Fraser RD, Vernon-Roberts B. Discitis after discography. The role of prophylactic antibiotics. *J.Bone Joint Surg.Br.* 1990;72:271-4.

16. Guyer RD, Ohnmeiss DD. Lumbar discography. Position statement from the North American Spine Society Diagnostic and Therapeutic Committee Spine 1995;20:2048-59.
17. Walsh TR, Weinstein JN, Spratt KF et al. Lumbar discography in normal subjects. A controlled, prospective study. J.Bone Joint Surg.Am. 1990;72:1081-8.
18. Merskey H, Bogduk N. Classification of chronic pain syndromes and definition of pain terms. Seattle: WAIASP Press, 1994:180-181.
19. Schwarzer AC, Aprill CN, Derby R et al. The prevalence and clinical features of internal disc disruption in patients with chronic low back pain. Spine 1995;20:1878-83.
20. Wetzel FT, LaRocca SH, Lowery GL et al. The treatment of lumbar spinal pain syndromes diagnosed by discography. Lumbar arthrodesis. Spine 1994;19:792-800.
21. Olerud S, Karlstrom G, Sjostrom L. Transpedicular fixation of thoracolumbar vertebral fractures Clin.Orthop.Relat Res. 1988;227:44-51.
22. Esses SI, Botsford DJ, Kostuik JP. The role of external spinal skeletal fixation in the assessment of low-back disorders. Spine 1989;14:594-601.
23. van der Schaaf DB, van LJ, Pavlov PW. Temporary external transpedicular fixation of the lumbosacral spine. Spine 1999;24:481-4.
24. McCulloch JA, Waddell G. Lateral lumbar discography. Br.J.Radiol. 1978;51:498-502.
25. Fraser RD, Osti OL, Vernon-Roberts B. Discitis after discography. J.Bone Joint Surg.Br. 1987;69:26-35.
26. Adams MA, Dolan P, Hutton WC. The stages of disc degeneration as revealed by discograms. J.Bone Joint Surg.Br. 1986;68:36-41.
27. Willems PC, Jacobs W, Duinkerke ES et al. Lumbar discography: should we use prophylactic antibiotics? A study of 435 consecutive discograms and a systematic review of the literature. J.Spinal Disord.Tech. 2004;17:243-7.
28. Fritzell P, Hagg O, Wessberg P et al. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. Spine 2001;26:2521-32.
29. Dawson EG, Lotysch M, III, Urist MR. Intertransverse process lumbar arthrodesis with autogenous bone graft. Clin.Orthop.Relat Res. 1981;90-6.
30. Jackson RK, Boston DA, Edge AJ. Lateral mass fusion. A prospective study of a consecutive series with long-term follow-up. Spine 1985;10:828-32.
31. Lee CK, Langrana NA. Lumbosacral spinal fusion. A biomechanical study. Spine 1984;9:574-81.

32. Lehmann TR, Spratt KF, Tozzi JE et al. Long-term follow-up of lower lumbar fusion patients. *Spine* 1987;12:97-104.
33. Sato H, Kikuchi S. The natural history of radiographic instability of the lumbar spine. *Spine* 1993;18:2075-9.
34. Kumar MN, Jacquot F, Hall H. Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease. *Eur.Spine J.* 2001;10:309-13.
35. Throckmorton TW, Hilibrand AS, Mencia GA et al. The impact of adjacent level disc degeneration on health status outcomes following lumbar fusion. *Spine* 2003;28:2546-50.
36. Hambly MF, Wiltse LL, Raghavan N et al. The transition zone above a lumbosacral fusion. *Spine* 1998;23:1785-92.
37. Modic MT, Steinberg PM, Ross JS et al. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *Radiology* 1988;166:193-9.
38. van Tulder MW, Assendelft WJ, Koes BW et al. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine* 1997;22:427-34.
39. Gillet P. The fate of the adjacent motion segments after lumbar fusion. *J.Spinal Disord.Tech.* 2003;16:338-45.
40. Nachemson A. Lumbar discography--where are we today? *Spine* 1989;14:555-7.
41. Carragee EJ, Tanner CM, Khurana S et al. The rates of false-positive lumbar discography in select patients without low back symptoms. *Spine* 2000;25:1373-80.
42. Soini JR, Seitsalo SK. The external fixation test of the lumbar spine. 30 complications in 25 of 100 consecutive patients. *Acta Orthop.Scand.* 1993;64:147-9.
43. Jeanneret B, Jovanovic M, Magerl F. Percutaneous diagnostic stabilization for low back pain. Correlation with results after fusion operations. *Clin.Orthop.Relat Res.* 1994;130-8.
44. Lund T, Oxland TR, Nydegger T et al. Is there a connection between the clinical response after an external fixation test or a subsequent lumbar fusion and the pre-test intervertebral kinematics? *Spine* 2002;27:2726-33.
45. Zanoli G, Stromqvist B, Jonsson B. Visual analog scales for interpretation of back and leg pain intensity in patients operated for degenerative lumbar spine disorders. *Spine* 2001;26:2375-80.
46. Price DD, McGrath PA, Rafii A et al. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain* 1983;17:45-56.

47. Melzack R, Katz J. Pain measurement in persons in pain. In: Wall PD, Melzack R, eds. Text book of pain. Edinburg: Churchill Livingstone, 1994:337-51.
48. Hagg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur.Spine J.* 2003;12:12-20.

4b

Lumbar discography: Should we use prophylactic antibiotics? A study of 435 consecutive discograms and a systematic review of the literature

Willems PC, Jacobs WCH, Duinkerke ES, de Kleuver M

J Spinal Disord Tech 2004;17:243-247

ABSTRACT

Background. Lumbar discography can be used in the diagnostic work-up of degenerative spine disease. The most serious complication is discitis, believed to be due to penetration of the disc by a needle contaminated with skin flora. The use of prophylactic antibiotics has been advocated, although there is great concern regarding their efficacy and possible adverse effects on disc cells.

Methods. In the current study, the incidence of postdiscography discitis without the use of prophylactic antibiotics was studied in a consecutive patient group. Additionally, a systematic literature review was performed using strict criteria: 1) Discography was performed by means of a two-needle technique, 2) complications such as discitis were specifically looked for at follow-up, and 3) the exact numbers of patients and those lost to follow-up were reported.

Results. The clinical results of 200 patients with 100% follow-up for a minimum period of 3 months showed no case of discitis. In the literature review, 10 studies were selected. Nine studies without prophylactic antibiotics reported an overall incidence of 12 cases in 4891 patients (0.25%) or 12,770 discs (0.094%). The only study with prophylactic antibiotics (127 patients) showed no case of discitis.

Conclusions. Regarding the small number of patients in the only study in which antibiotics were used and the overall low incidence of postdiscography discitis, not enough evidence was found that prophylactic antibiotics can prevent discitis. It was concluded that in lumbar discography by means of a two-needle technique without prophylactic antibiotics, the risk of postdiscography discitis is minimal and there is not enough support from the literature to justify the routine use of prophylactic antibiotics.

INTRODUCTION

As part of the diagnostic work-up in the evaluation of degenerative spine disease, lumbar discography can be performed in case the clinical problem may require surgical treatment^{1,2}. Next to visualizing the internal structure of the disc, the benefit of discography is its ability to distinguish between symptomatic and asymptomatic degenerative discs. In this way, when planning an arthrodesis of the spine, the number of motion segments to be fused can be determined.

Discography is an invasive procedure with a fairly low complication rate³⁻⁵. The most serious complication is discitis, believed to be due to penetration of bacteria into the disc by a needle contaminated with skin flora⁶. Although very few authors seem to have specifically checked for postdiscography discitis at follow-up, generally low incidences have been reported³. As discitis is a serious complication for a diagnostic procedure, prevention by the use of prophylactic antibiotics, either intravenously or intradisally, has been advocated⁷. There is, however, controversy concerning the ability of intravenous antibiotics to penetrate the disc and reach a therapeutic level^{8,9}. Furthermore, little is known about possible side effects of the use of intradiscal antibiotics. As damaging effects on disc tissue have already been reported,¹⁰ caution regarding the use of prophylactic antibiotics is warranted.

In search for a clinical guideline in lumbar discography, the aims of the current study were 1) to determine clinically postdiscography discitis incidence in a consecutive group of patients without the use of prophylactic antibiotics and 2) to perform a systematic literature review to study postdiscography discitis incidence in general.

MATERIALS AND METHODS

During 1 year, the records of all consecutive patients undergoing lumbar discography as part of the work-up for degenerative lumbar spine disease were investigated. All procedures were carried out in the Radiology Department by orthopaedic residents affiliated with the spine unit. No standard preoperative sedation was used, and no prophylactic antibiotics were administered. According to the technique described by McCulloch and Waddell,³ the patients were placed in the left lateral decubitus

position. The skin was prepared with Betadine and draped in a routine manner. An image intensifier was used to provide both lateral and anteroposterior views of the spine. After superficial infiltration of local anesthetic, a solid stilette double needle (18-gauge outer and 22-gauge inner solid needle) was advanced to the intended disc space. After fluoroscopic confirmation that the needle was in the center of the disc, the inner solid needle was removed, and Isovist, a watersoluble, nonionic contrast dye (Schering AG, Berlin, Germany), was injected slowly. If the disc center could not be approached properly, an open, more flexible 22-gauge needle was advanced through the outer needle to reach the center. Volume and pressure of the dye injected as well as its distribution pattern in the disc were scored according to the criteria of Adams et al,¹¹ and the patient's response was registered. For each disc level, separate needles were used. One hour after the procedure, patients were sent home with the instruction to return immediately in case of aggravated back pain or fever. Routine control followed at 6-12 weeks after discography. In case of pain, discomfort, or fever, laboratory and radiologic controls were performed. Discitis was diagnosed when erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) were raised and magnetic resonance imaging (MRI) revealed the typical increased bony signal in the vertebral bodies with destruction of the endplates and paravertebral signal enhancement¹². In those cases where the records did not reveal the postdiscography status or follow-up had been <12 weeks, patients were approached for a telephone interview and questioned for the above-mentioned symptoms during the first 3 months after the procedure.

Literature Search Strategy and Selection Criteria

To obtain all relevant literature, the most common databases of published literature were used:

Medline (1966 to January 2003)

Current Contents (1996 to February 2003)

Cochrane database of randomized controlled trials (2003, issue 1)

The search strings with the number of matches for each string are shown in Table 1. No restrictions were made on language or date. On the basis of the abstracts, articles were selected by two independent reviewers. The references from the selected articles were checked in the same way to ensure no relevant articles were missed.

Table 1. Number of Articles Found with Search Strings and Databases Used

Search Strings	Database		
	Medline	Current Contents	Cochrane
Diagnosis			
Discitis (mesh)	690	101	3
Discitis (tw)	875	170	9
Infection (tw)	403,443	183,636	17,246
Postoperative complications (mesh)	244,608	893	13,919
Complications (tw)	1,288,535	82,472	37,381
Union diagnoses	1,625,657	256,529	54,355
Treatment			
Intervertebral disc/radiography* (mesh)	1182	357	60
Discography (tw)	439	597	18
Union treatment	1390	935	73
Indication			
Lumbar (tw)	44,472	14,341	2615
Lumbar vertebrae (mesh)	19,572	132	647
Union study design	44,472	14,341	2615
Intersection	219	36	24
Intersection databases			263

Articles were finally included when they met the criteria in the following checklist:

- Lumbar discography should be performed at all levels with an adequately described two-needle technique under strict aseptic conditions.
- The presence of complications such as postdiscography discitis must have been specifically looked for at follow-up and should be noted as such.
- The exact number of patients and those lost to follow-up should be noted.
- Preferably a minimum follow-up of 3 months was used. If both reviewers could not reach consensus, a third reviewer was consulted.

RESULTS

Discography was performed in exactly 200 patients (88 men, 112 women) with a mean age of 41 (16-62) years. Forty patients had discography at one level, 85 patients at two levels, and in 75 patients three levels were involved. Thus, in total 435 discs were injected. Three patients had reported signs of discomfort or fever at follow-up visit. Laboratory controls, however, showed no significant elevation of ESR or CRP and their MRI scan was negative for discitis. As their symptoms resolved spontaneously, discitis was excluded. As all patients could be reached by phone, 100% follow-up was guaranteed. No patient reported complaints of fever or discomfort regarding his or her back in the first 3 months after discography, and thus an overall incidence of 0% postdiscography discitis was registered.

Table 2. Discitis Incidences in Selected Literature and Current Study by Patient and by Disc

Study	Year	Prophylactic Antibiotics	No. of Patients
Collis and Gardner ¹³	1962	No	1014
Patrick ¹⁴	1973	No	123
Colquhoun ¹⁵	1977	No	716
Milette and Melanson ¹⁶	1982	No	500
Fraser et al ⁶	1987	No	210
Guyer et al ⁴	1988	No	2014
Jackson et al ¹⁷	1989	No	124
Simmons et al ¹⁸	1991	No	164
Carragee et al ¹⁹	2000	No	26
Current study	2003	No	200
Literature combined with current study	—	No	5091
Osti et al ⁷	1990	Yes	127

Literature Search

The search yielded 263 references, of which 217 papers were excluded on the basis of an abstract and title that showed no relation to the human lumbar spine in vivo and discography. The remaining 46 papers were subjected to a thorough investigation with the aid of the above-mentioned criteria. Additionally, the references were screened according to the same rules to make sure that no relevant articles were missed. Finally, 10 papers,^{4,6,7,13-19} spanning a period of 40 years, were selected and listed in Table 2. Two of these studies^{7,19} actually reported a minimum follow-up of 3 months, whereas the remaining studies did not specify the exact term of follow-up. In only one of the selected studies had prophylactic antibiotics been used, and a postdiscography discitis incidence of 0% was found.⁷ In the other nine studies, no prophylactic antibiotics had been administered. These studies revealed 12 cases of discitis in 4891 patients and 12,770 discs with an incidence of

No. of Discs	Discitis Cases	Incidence by Patient (%)	Incidence by Disc (%)
2187	1	0.10	0.05
341	0	0.00	0.00
2000	1	0.14	0.05
1009	3	0.60	0.30
417	4	1.90	0.98
6042	2	0.10	0.05
231	0	0.00	0.00
465	1	0.61	0.22
78	0	0.00	0.00
435	0	0.00	0.00
13,205	12	0.24	0.09
337	0	0.00	0.00

0.25% by patient and 0.094% by disc. Three of these studies,^{14,17,19} as well as the current study, reported no case of discitis.

Combining the literature results with the current patient study leads to an overall incidence of 0.24% by patient and 0.091% by disc.

DISCUSSION

Discitis after discography has been recognized as a distinct entity that is considered to be due to bacterial infection^{6,20} and should not be confused with the common exacerbation of backache after intradiscal injection most probably due to a mechanical or chemical cause. The most common causative agents are considered to be *Staphylococcus aureus* and *S. epidermidis*, although the incidence of gram-negative bacteria appears to be rising^{10,21,22}. Fraser et al⁶ isolated bacteria from the disc in three of four discitis patients who had open biopsy <6 weeks after discography, and they suggested bacteria introduced by the needle tip to be the initiators of the disease. To eliminate the risk of injecting a core of soft tissue including skin, the use of stilette needles and a two-needle technique was recommended, with which the incidence of discitis could be reduced from 2.7% to 0.7%. In a prospective clinical trial of 127 patients undergoing lumbar discography with cephazolin added to the contrast dye, Osti et al⁷ reported no case of discitis and recommended the use of a single dose of prophylactic broad-spectrum antibiotic either added to the radiographic contrast material or administered intravenously. However, Eismont et al²³ reported no detectable levels in discs of rabbits after intravenous administration of either cephalothin or oxacillin, and Gibson et al⁸ was unable to retrieve any cephradine or flucloxacillin from human intervertebral discs intraoperatively in 25 scoliosis patients. Boscardin et al⁹ found therapeutic levels in only a short period after intravenous administration of high doses of cefazolin, whereas Rhoten et al²⁴ found quantifiable levels of cefazolin and oxacillin in only 40% of human cervical discs. As the disc is an avascular structure with blood vessels found only at its margins,²⁵ intradiscal nutrition and levels of permeation of drugs depend on passive diffusion through the adjacent endplates and the surrounding annulus fibrosis. The disc has a negatively charged structure, which implies that positively charged antibiotics such as gentamicin may diffuse

more rapidly and be more effective than the negatively charged antibiotics such as cephalosporins²⁶. In a prospective study of 20 patients undergoing spinal surgery, Tai et al²¹ could detect no therapeutic level of cefuroxime in the disc, whereas gentamicin showed a 50% penetration from blood to disc, well within the therapeutic range. They suggested, in accordance with Rhoten et al²⁴ and Langmayr et al,²⁷ that high doses of cephalosporins are required to achieve therapeutic levels in the disc. Important in this respect are the recent findings by Hoelscher et al¹⁰ that high doses of antibiotics, including both cephalosporins and aminoglycosides, can have deleterious effects on the survival of cultured disc cells, cell proliferation, and metabolic rates. This may have serious implications if antibiotics are to be mixed with contrast dye and injected directly into the disc, creating local peak concentrations. Moreover, side effects of aminoglycosides are well known, and sensitization or anaphylactic reactions should always be anticipated²⁸.

In the current study, lumbar discography was performed in 200 consecutive patients by means of a stilette two-needle technique under strict aseptic conditions without the use of prophylactic antibiotics. A discitis incidence of 0% was found. As method of study, the routine postclinical assessment was used, which we consider crucial in the detection of postdiscography discitis. It has been stated that underestimation of discitis incidence is possible because of the latent period between discography and the onset of symptoms, lack of clinical contact between patient and clinician, and lack of awareness by the clinician^{6,29}. Therefore, a routine follow-up visit is mandatory. As the onset of symptoms of discitis after uneventful discography normally develops within a few weeks and at least within 10-12 weeks,^{29,30} we believe no case of discitis has gone undetected by our method.

In a review of 15 studies, the North American Spine Society⁵ calculated an overall incidence of <0.15% by patient and <0.08% by disc. Not all of these studies reported the exact number of patients and those lost to follow-up, and no distinction was made between studies using a one- or a two-needle technique. In the current literature search, without the use of prophylactic antibiotics, overall incidences of 0.25% by patient and 0.094% by disc were found (see Table 2). In one of these selected papers,⁶ unusually high incidences (1.90% and 0.98%,

respectively) in 210 patients were reported. This might have been caused by the fact that in 61 of these patients, discography was performed by less experienced physicians, leading to 3 cases of discitis (4.9%), as opposed to 1 case in 149 patients (0.7%) if the procedure had been carried out by an experienced discographer. In the current literature search, only one study of lumbar discography with the use of prophylactic antibiotics, in which discitis was specifically was looked for, was selected⁷. In this series of 127 patients (337 discs), no case of discitis was registered. However, to prove that additive antibiotics actually prevent postdiscography discitis, given the incidence of 0.25% by patient, a randomized trial of approximately 9000 patients would be needed to reach significance (calculated sample size for dichotomous response variables). Taking into account the possible adverse effects of antibiotics, we feel there is not enough evidence to regard the routine use of prophylactic antibiotics indicated in lumbar discography. At present, as has been stressed before,³¹ it is considered imperative that discography be carried out by a physician who is very experienced with the procedure, performs it on a routine basis, and is meticulous with the two-needle technique and aseptic conditions.

CONCLUSIONS

This study has shown that lumbar discography performed with a stilette two-needle technique without prophylactic antibiotics can be a safe procedure with a minimal risk of development of discitis. As there is no evidence in the literature that prophylactic antibiotics actually reduce the incidence of discitis, we feel that the routine use of prophylactic antibiotics in lumbar discography is not indicated.

REFERENCES

1. Lindblom K. Diagnostic puncture of intervertebral disks in sciatica. *Acta Orthop Scand*. 1948;17:231-239.
2. Greenspan A, Amparo EG, Gorczyca DP, et al. Is there a role for diskography in the era of magnetic resonance imaging? Prospective correlation and quantitative analysis of computed tomography-diskography, magnetic resonance imaging, and surgical findings. *J Spinal Disord*. 1992;5: 26-31.
3. McCulloch JA, Waddell G. Lateral lumbar discography. *Br J Radiol*. 1978;51:498-502.
4. Guyer RD, Collier R, Stith WJ, et al. Discitis after discography. *Spine*. 1988;13:1352-1354.
5. Guyer RD, Ohnmeiss DD. Lumbar discography. Position statement from the North American Spine Society Diagnostic and Therapeutic Committee. *Spine*. 1995;20:2048-2059.
6. Fraser RD, Osti OL, Vernon-Roberts B. Discitis after discography. *J Bone Joint Surg (Br)*. 1987;69:26-35.
7. Osti OL, Fraser RD, Vernon-Roberts B. Discitis after discography. The role of prophylactic antibiotics. *J Bone Joint Surg (Br)*. 1990;72:271-274.
8. Gibson MJ, Karpinski MR, Slack RC, et al. The penetration of antibiotics into the normal intervertebral disc. *J Bone Joint Surg (Br)*. 1987;69:784- 786.
9. Boscardin JB, Ringus JC, Feingold DJ, et al. Human intradiscal levels with cefazolin. *Spine*. 1992;17:S145-S148.
10. Hoelscher GL, Gruber HE, Coldham G, et al. Effects of very high antibiotic concentrations on human intervertebral disc cell proliferation, viability, and metabolism in vitro. *Spine*. 2000;25:1871-1877.
11. Adams MA, Dolan P, Hutton WC. The stages of disc degeneration as revealed by discograms. *J Bone Joint Surg (Br)*. 1986;68:36-41.
12. Schulitz KP, Assheuer J. Discitis after procedures on the intervertebral disc. *Spine*. 1994;19:1172-1177.
13. Collis JSJ, Gardner WJ. Lumbar discography: An analysis of one thousand cases. *J Neurosurg*. 1962;19:452-461.
14. Patrick BS. Lumbar discography: a five year study. *Surg Neurol*. 1973;1: 267-273.
15. Colquhoun J. The role of lumbar discography in diagnosis of lumbar spine lesions. *Orthop Rev*. 1977;6:51-56.
16. Milette PC, Melanson D A reappraisal of lumbar discography. *J Can Assoc Radiol*. 1982;33:176-182.

17. Jackson RP, Becker GJ, Jacobs RR, et al. The neuroradiographic diagnosis of lumbar herniated nucleus pulposus: I. A comparison of computed tomography (CT), myelography, CT-myelography, discography, and CTdiscography. *Spine*. 1989;14:1356-1361.
18. Simmons JW, Emery SF, McMillin JN, et al. Awake discography. A comparison study with magnetic resonance imaging. *Spine*. 1991;16: S216-S221.
19. Carragee EJ, Chen Y, Tanner CM, et al. Can discography cause long-term back symptoms in previously asymptomatic subjects? *Spine*. 2000;25: 1803-1808.
20. Fraser RD, Osti OL, Vernon-Roberts B. Discitis following chemonucleolysis. An experimental study. *Spine*. 1986;11:679-687.
21. Tai CC, Want S, Quraishi NA, et al. Antibiotic prophylaxis in surgery of the intervertebral disc. A comparison between gentamicin and cefuroxime. *J Bone Joint Surg (Br)*. 2002;84:1036-1039.
22. Rimoldi RL, Haye W. The use of antibiotics for wound prophylaxis in spinal surgery. *Orthop Clin North Am*. 1996;27:47-52.
23. Eismont FJ, Wiesel SW, Brighton CT, et al. Antibiotic penetration into rabbit nucleus pulposus. *Spine*. 1987;12:254-256.
24. Rhoten RL, Murphy MA, Kalfas IH, et al. Antibiotic penetration into cervical discs. *Neurosurgery*. 1995;37:418-421.
25. Urban JP, Holm S, Maroudas A et al. Nutrition of the intervertebral disk. An in vivo study of solute transport. *Clin Orthop*. 1977;129:101-114.
26. Thomas RW, Batten JJ, Want S, et al. A new in-vitro model to investigate antibiotic penetration of the intervertebral disc. *J Bone Joint Surg (Br)*. 1995;77:967-970.
27. Langmayr JJ, Obwegeser AA, Schwarz AB, et al. Intrathecal steroids to reduce pain after lumbar disc surgery: a double-blind, placebo-controlled prospective study. *Pain*. 1995;62:357-361.
28. Aprill CN. Diagnostic lumbar injections. In: Frymoyer JW, ed. *The Adult Spine: Principles and Practice*. New York: Raven Press; 1991:420-442.
29. Rohde V, Meyer B, Schaller C, et al. Spondylodiscitis after lumbar discectomy: incidence and a proposal for prophylaxis. *Spine*. 1998;23: 615-620.
30. Piotrowski WP, Krombholz MA, Muhl B. Spondylodiscitis after lumbar disk surgery. *Neurosurg Rev*. 1994;17:189-193.
31. Fraser RD. The North American Spine Society (NASS) on lumbar discography. *Spine*. 1996;21:1274-1276.

5

Temporary external transpedicular fixation of the lumbosacral spine A prospective, longitudinal study in 330 patients

Elmans L, Willems PC, Anderson PG, van Limbeek J, de Kleuver M, van der Schaaf DB

Spine 2005;30:2813-2816

ABSTRACT

Study Design. In this study, 330 patients with incapacitating low back pain underwent temporary external transpedicular fixation (TETF) of the lumbosacral spine in a prospective trial.

Objective. To evaluate TETF as a test for selecting suitable candidates for segmental spinal fusion.

Summary of Background Data. Few studies regarding TETF have been published, and contradictory results concerning predictive value and morbidity were reported.

Methods. All patients were tested with the external fixator in two different positions: fixation and nonfixation. Before and during the test and at follow-up examination, pain was assessed on a Visual Analogue Scale (VAS). The TETF test was considered to be positive if the VAS score in the fixation state was 30 or more points lower than in the nonfixation state. Hence, a positive test would imply the decision to perform segmental lumbosacral fusion. When the reduction was less than 30 points, the test was negative. Individual pain reduction and working capacity were taken as measure of outcome.

Results. Most of the patients in this study (62%) underwent spinal surgery previously. The positive and negative TETF groups were quite similar, but a large withingroup variation was found. Within the fusion group of 123 patients, improvement in VAS scores and improvement in working capacity were not significantly better for the positive TETF group in comparison with the negative TETF group.

Conclusion. In this heterogeneous group of chronic patients with low back pain, TETF of the spine (including a placebo trial) does not appear to be of value in selecting suitable candidates for spinal fusion.

INTRODUCTION

Surgical treatment for degenerative disc disease (DDD) of the lumbar spine is still controversial¹⁻⁴. Despite preoperative analysis by means of immobilization in a pantaloons plaster cast, provocative discography, facet blocks, and psychological screening, the result of lumbar fusion in chronic patients with low back pain is hard to predict, especially after previous lumbar surgery⁵⁻¹⁰. A reliable test that can predict whether a patient may benefit from fusion is needed.

Temporary external transpedicular fixation (TETF) of the lumbosacral spine has been used as a diagnostic tool for decision-making in low back pain. Results of mechanical studies suggest that external fixation constructs of the spine in a neutral (i.e., not compressed nor distracted) position provide a high degree of stabilization and that it is possible to relieve symptoms resulting from segmental spinal instability by using this construct¹¹⁻¹⁵.

Since the introduction of TETF by Olerud et al,¹⁶ some surgeons use the test to predict the clinical result of spinal fusion¹⁷⁻²⁰. In the later modified "dynamized state," the fixator bars are also fixed horizontally for some time, thereby providing no immobilization at the instrumented level²¹. With this placebo trial added, the patient serves as his or her own control. At this moment, there are no reports on the middle-long and long-term success of TETF with a placebo trial to predict fusion outcome.

The aim of this long-term prospective study in patients with low back pain is to evaluate whether TETF, including a placebo trial, can predict the clinical outcome of lumbar segmental fusion.

METHODS

Patients with incapacitating low back pain for at least 1 year were selected for the test. Before TETF, the patients had undergone conventional diagnostic tests, such as radiologic and/or MRI imaging of the lumbar spine, and a pantaloons plaster cast testing or provocative discography. On the basis of the data acquired with these examinations, the symptomatic levels to be tested were determined. Only when it remained unclear after these diagnostic tests whether the patient would benefit from a fusion was the patient included in this study.

During the TETF test, patients were hospitalized for 7 to 10 days. The insertion of the external fixator was performed under general anesthesia with fluoroscopic control. Initially, no antibiotics were given; however, since May 1997, antibiotic prophylaxis (oral flucloxacillin 500 mg QID) was administered before the operation and continued until discharge from the hospital. The AO tubular external fixator with self-tapping 5-mm diameter Schanz's screws was used. The external fixator was tested in two positions: fixation and nonfixation (dynamisation). In the latter, the connecting bars were placed horizontally, connecting two Schanz screws of the same vertebra. Hence, there was no immobilization of the instrumented levels. For both positions, pain on a Visual Analogue Scale (VAS) was assessed. Each test position was maintained for 2 days. The patient was unaware whether the instrumented levels were fixed or dynamized. The surgeon randomly chose order of adjustment of the fixation. The first adjustment was only made after 3 days to give the patient time to get used to the device and to provide time to overcome postoperative pain. At the end of the test period, the screws were removed under general anesthesia.

Suitable candidates for fusion were those patients who experienced adequate pain relief in fixation and definitely less pain relief in the nonfixation position.

At the time of admission and before the test, pain on a VAS ranging from 0 (no pain) to 100 points (maximal pain), subjective pain sensations (good, reasonable, fair, poor), and working capacity were noted. Working capacity was determined by asking the patients to estimate what percentage of their jobs they were able to do. For those receiving compensation, subtracting the percentage of compensation from 100 made the estimation. Follow-up assessment, performed with a written questionnaire, evaluated the same parameters. The TETF test result, type of treatment after the test (fusion or conservative treatment and the type of fusion), age, gender, duration of symptoms before the test, and preoperative diagnosis were the variables tested.

Between October 1989 and October 2000, 330 patients (207 women and 123 men, average age 42 years, SD = ± 9 years) underwent TETF. Patients had a wide variety of preoperative diagnoses with the majority of patients (204 of 330 cases) having undergone spinal surgery (Table 1). The mean duration of their symptoms was 6 years (SD = ± 5 years). In the majority of patients, L4-S1 (n = 150; 44%) or L5-S1 (n = 78; 23%) was tested.

Table 1. Preoperative Diagnosis of 330 Patients Undergoing TETF

	No. (%)
Previously operated	205 (62)
Postlaminectomy syndrome	69 (21)
Pain after herniotomy or chemonucleolysis	68 (21)
Pain after (solid) fusion (adjacent segment)	21 (6)
Pain after (doubtful) fusion	20 (6)
Pseudarthrosis	27 (8)
Not previously operated	125 (38)
Degenerative disc disease	100 (30)
Spondylolysis/listhesis	20 (6)
Other (post-traumatic/discitis)	5 (2)
Total	330 (100)

The median follow-up period was 79 months (range, 15-144 months). Patients had to complete the test to be included in the study. Two patients did not complete the TETF due to complications. Questionnaire data are missing for another 27 patients (2 deceased, 8 patients could not be found, 17 forms were not returned in spite of repeated reminders). Thus, the response was 92% consisting of 108 male and 193 female patients. Unless otherwise specified, 301 subjects have been used for the calculations. No differences in age, gender, duration of symptoms, or length of follow-up were found between the patients who returned the questionnaire and those who did not. For 272 patients (272 of 328 = 83% returned questionnaires), working capacity both before and after the test could be determined: in addition to the abovementioned 27 patients, 29 of 301 patients who completed the questionnaire did not report their current working capacity.

Test Parameter. The TETF evaluation parameter was the VAS score at fixation subtracted from the VAS score at nonfixation. A reduction of 30 points or more in comparison to the nonfixation VAS score was defined as a positive TETF test. When the reduction was less than 30 points, the test was defined to be negative.

Outcome Parameters. The VAS score at follow-up subtracted from the VAS score before the test was used as the outcome parameter. A reduction of 30% or more in comparison to the initial VAS score was considered to be a clinically relevant improvement in pain. Since subjective results correlated with change in VAS score ($r = 0.65$, $P = 0.00$), only the latter was used as outcome parameter for pain. An additional outcome parameter was the change in working capacity at follow-up subtracted from the working capacity before the test.

Statistics. Nominal- and ordinal-scaled variables were analyzed using χ^2 tests. Differences in means of continuous and interval scaled variables (age and duration of symptoms) were analyzed by paired and unpaired Student's t test and analyses of variance, with alpha set at 0.05. Pearson Moment Correlation Coefficients estimated associations.

RESULTS

Change in VAS score correlated with subjective results ($r = 0.65$, $P = 0.00$) and was used as outcome parameter for pain in the way as described in Methods (including calculations with 30% differences).

The main complications associated with TETF were pin track infection and nerve root irritation. No patients experienced leakage of cerebrospinal fluid. Pin track infection occurred in 20 of 330 patients (6%). Surgical treatment to cure the infection was required in 2 patients (0.6%). The rest healed uneventfully after antibiotic treatment and removal of the external fixator at completion of the test. Twenty-six of 330 patients (8%) had nerve root irritation that recovered after the test in these patients without signs of permanent neural damage. Two patients with root irritation had so much pain during the test that proper evaluation was impossible and the external fixator was removed before completion of the test. A positive TETF test result was found in 125 patients, whereas a negative test result was found in 176 patients. The statistical analysis revealed a weak (0.51 , $P = 0.00$) correlation in TETF test result and the type of treatment; 70% (88 of 125) of the patients with a positive TETF test result were treated by fusion, whereas 80% (141 of 176) of the patients with a negative TETF score were not operated on. The

groups were quite similar but a large within group variation was found (Table 2).

The mean follow-up VAS scores were on average 8 points lower (on the basis of 100) in the group of patients who had a positive TETF test result (with or without operation following) relative to the group of patients with a negative test (Table 3). This difference is not of clinical relevance and not statistically significant. Improvement in working capacity was not significantly related to TETF test result (Table 3).

Table 2. Descriptives for Groups With Negative and Positive TETF Result in 301 Responders

	Negative	Positive
Age (yr)	41.8 ± 9.3	40.7 ± 8.7
Duration of complaints (yr)	6.1 ± 6.7	5.5 ± 5.1
Previous operations done (%)	64.8	61.6
Degenerative disc disease (%)	30.7	32.0
Failed back surgery cases (%)	61.9	60.0
Lysis/listhesis (%)	6.4	6.3
Other diagnosis (%)	1.2	4.0
Length of follow-up (mo)	80.8 ± 34.6	75.6 ± 33.8
Total no. of patients = 301	176	125

Table 3. Mean VAS Score and Working Capacity in Groups With Negative and Positive TETF Result

	Negative	Positive
VAS score		
Before TETF	73.6 ± 15.1	73.8 ± 17.7
Follow-up	62.9 ± 26.8	54.0 ± 31.7
Difference (before TETF — follow-up)	10.7 ± 28.3	19.8 ± 32.2
Total no. of patients = 301	176	125
Working capacity		
Before TETF (%)	19.7 ± 30.2	21.8 ± 31.8
Follow-up (%)	30.6 ± 33.0	38.5 ± 37.1
Difference (before TETF — follow-up)	10.9 ± 37.0	16.7 ± 41.0
Total no. of patients = 272	160	112

Of the 301 patients, 123 (41%) underwent a definite fusion operation within 3 to 6 months (fusion group), and 178 (59%) were treated without operation (nonoperative group). The median follow-up period was 78 months (range, 18-144 months) for the fusion group and 76 months (range, 15-144 months) for the nonoperative group. Seventy-three percent (90 of 123) of operations were performed by posterolateral fusion supplemented with pedicle screw fixation and 33 of 123 (27%) via an anterior procedure. No differences in result between these groups were found.

Within the fusion group, patients with a negative TETF test result showed a mean VAS score improvement of 11 (± 33) points at follow-up, whereas the group of patients with a positive TETF test result showed a mean improvement of 22 (± 31) points (Table 3). This 11- point difference is not statistically significant because of the large within-group variation. Improvement in working capacity was not significantly better in the positive TETF group than in the negative TETF group (Table 4).

Table 4. Fusion Group: VAS Score sand Working Capacity in Groups With Negative and Positive TETF Result and Results of T Test for Equality of Means

	Negative	Positive	df	P
VAS scores				
Total no. of patients = 123	35	88		
Difference (before TETF — follow-up) (%)	11.2 \pm 33.5	22.2 \pm 30.6	121	0.08
Working capacity				
Total no. of patients = 108	31	77		
Difference (before TETF — follow-up) (%)	13.2 \pm 35.3	15.4 \pm 38.0	106	0.77

Similar results are found if one compares improvement in the fusion group. Within the patients with a negative TETF, 29% showed an improvement in the VAS score while 46% of patients with a positive TETF showed improvement in VAS (Table 5). This 17% difference is not statistically significant, again probably because of the same large within-group variation.

Table 5. Fusion Group: Negative and Positive TETF Result vs. Percentage Improvement of >30% or <30% in VAS Score

	Category % Improvement (VAS score before TETF — follow-up)		Total
	<30%	≥30%	
Negative TETF (no.)	25	10	35
% within negative TETF	71	29	100%
Positive TETF (no.)	48	40	88
% within positive TETF	55	45	100%
Total (no.)	73	50	123

DISCUSSION

In a previous study with the earlier results of part of the current patient group,¹⁹ TETF appeared to be a valuable tool for surgical decision-making in patients with low back pain. The present study with a much larger patient group and longer follow-up does not confirm those findings.

Patients with a positive and negative TETF test result were equivalent in age, gender, duration of symptoms, and preoperative diagnosis. Between positive and negative TETF test groups, no clinically relevant improvement in VAS score or working capacity was found at follow-up. Within the fusion group, pain reduction did not differ between patients with a positive and those with a negative TETF test.

Only a weak correlation ($r = 0.5$) between the TETF test result and the type of treatment was found. After the TETF test, the final decision whether to operate or not was based on result of the test combined with the clinical judgment of patient and surgeon, whereas statistical analysis at follow-up was performed by using the exact calculated TETF test result. This explains the fact that despite a positive TETF test in 30% of the patients an operation was not performed, whereas 20% of the patients with a negative TETF score were treated by fusion.

A serious drawback of the test is that it is an invasive procedure, involving the percutaneous placement of pedicle screws under fluoroscopic control. Misplacement

of the screws can lead to neurologic damage. The 8% nerve root irritation in the current study is in concordance with the 0% to 9% reported in other series of external skeletal fixations^{17,18,22-25}. In 2 patients, the external fixator had to be removed before completion of the test due to pain caused by root irritation. No permanent neurologic deficit was found. Pin track infection occurred in 6% of the patients, only 2 patients needing surgical treatment for this infection. Other authors reported an infection rate of 2% up to 18%^{17,18,22-25}.

Probably, complication rate increases with the duration of the test²³. Since May 1997, flucloxacillin was administered orally in this trial as long as the external fixator was in situ. This coincided with a drop in infection rate from 7% before to 0%, indicating that prophylactic use of antibiotics prevents infection in these patients.

Although we realize that pain is only one of many determining factors leading to the suffering of patients with low back pain, pain on a VAS was chosen for the assessment of outcome. It is impossible to use advanced back pain scores such as Oswestry at every position of the fixator bars, so for methodologic reasons a pain score had to be used. The change in VAS score at follow-up correlated reasonably well with subjective results. A large intrinsic error in registration of a VAS score can be found: "Putting a dot on a line" is inherently inaccurate. Nevertheless, it remains an objective method of outcome evaluation, as the patient serves as his own control. The lack of homogeneity in VAS scores was corrected by dichotomizing the scores in this analysis by defining the test result as positive when the difference in VAS scores in the fixation state subtracted from the nonfixation state was ≥ 30 and as negative when this difference was less than 30 points. Also, a reduction of $\geq 30\%$ at follow-up in comparison to the initial VAS score was considered to be a clinically relevant improvement in pain. When used in this dichotomized manner, the results are not only statistically but also clinically relevant.

CONCLUSION

The results of this study show that in this heterogeneous group of chronic patients with low back pain, of whom 62% underwent spinal surgery previously, TETF with a fixation and a nonfixation period fails to predict the long-term outcome of lumbar segmental fusion and that the procedure has complications. Therefore, TETF

appears to be of no value as a tool for surgical decision making in this heterogeneous group of patients with chronic low back pain.

REFERENCES

1. Hanley EN Jr, David SM. Current Concepts Review: lumbar arthrodesis for the treatment of low back pain. *J Bone Joint Surg Am* 1999;81:716-30.
2. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 1999;24:1820-32.
3. Fritzell P, Hägg O, Wessberg P, et al. 2001 Volvo award study in clinical studies: lumbar fusion versus nonsurgical treatment for chronic low back pain. *Spine* 2001;26:2521-34.
4. Krismer M. Fusion of the lumbar spine: a consideration of the indications. *J Bone Joint Surg Br* 2002;84:783-94.
5. Markwalder TM, Reulen HJ. Diagnostic approach in instability and irritative state of a 'lumbar motion segment' following disc surgery-failed back surgery syndrome. *Acta Neurochir (Wien)* 1989;99:51-7.
6. Rask B, Dall BE. Use of the pantaloone cast for the selection of fusion candidates in the treatment of chronic low back pain. *Clin Orthop* 1993;288:148-57.
7. Esses SI, Moro JK. The value of facet joint blocks in patient selection for lumbar fusion. *Spine* 1993;18:185-90.
8. Axelsson P, Johnsson R, Stromqvist B, et al. Orthosis as prognostic instrument in lumbar fusion: no predictive value in 50 cases followed prospectively. *J Spinal Disord* 1995;8:284-8.
9. Derby R, Howard MW, Grant JM, et al. The ability of pressure-controlled discography to predict surgical and nonsurgical outcomes. *Spine* 1999;24:364-71.
10. Trief PM, Grant W, Fredrickson B. A prospective study of psychological predictors of lumbar surgery outcome. *Spine* 2000;25:2616-21.
11. Bednar DA, Raducan V. External spinal skeletal fixation in the management of back pain. *Clin Orthop* 1996;322:131-45.
12. Axelsson P, Johnsson R, Stromqvist B. Mechanics of the external fixation test in the lumbar spine: a roentgen stereophotogrammetric analysis. *Spine* 1996;21:330-33.
13. Sahni IK, Hipp JA, Kirking BC, et al. Use of percutaneous transpedicular external fixation pins to measure intervertebral motion. *Spine* 1999;24:1890-3.
14. Lund T, Nydegger T, Rathonyi G, et al. Three-dimensional stabilization provided by the external spinal fixator compared to two internal fixation devices: a biomechanical in vitro flexibility study. *Eur Spine J* 2003;12:474-9.

15. Lund T, Oxland TR, Nydegger T, et al. Is there a connection between the clinical response after an external fixation test or a subsequent lumbar fusion and the test intervertebral kinematics? *Spine* 2002;27:2726-33.
16. Olerud S, Slostrom L, Karlsrom G, et al. Spontaneous effect of increased stability of the lower lumbar spine in cases of severe chronic back pain: the answer of an external transpedicular fixation test. *Clin Orthop* 1988;203:67-74.
17. Soini J, Slätis P, Kannisto M, et al. External transpedicular fixation test of the lumbar spine correlates with the outcome of subsequent lumbar fusion. *Clin Orthop* 1993;293:89-96.
18. Jeanneret B, Jovanovic M, Magerl F. Percutaneous diagnostic stabilization for low back pain: correlation with results after fusion operations. *Clin Orthop* 1994;304:130-8.
19. van der Schaaf DB, van Limbeek J, Pavlov P. Temporary external transpedicular fixation of the lumbosacral spine. *Spine* 1999;24:481-5.
20. Axelsson P, Johnsson R, Stromqvist B, et al. Temporary external pedicular fixation versus definitive bony fusion: a prospective comparative study on pain relief and function. *Eur Spine J* 2003;12:41-7.
21. Esses SI, Botsford DJ, Kostuik JP. The role of external spinal skeletal fixation in the assessment of low-back disorders. *Spine* 1989;14:594-601.
22. Magerl FP. Stabilization of the lower thoracic and lumbar spine with external skeletal fixation. *Clin Orthop* 1984;189:125-41.
23. Soini JR, Seitsalo SK. The external fixation test of the lumbar spine: 30 complications in 25 of 100 consecutive patients. *Acta Orthop Scand* 1993; 64:147-9.
24. Faraj AA, Akasha K, Mulholland RC. Temporary external fixation for low back pain: is it worth doing? *Eur Spine J* 1997;6:187-90.
25. Bednar DA. Failure of external spinal skeletal fixation to improve predictability of lumbar arthrodesis. *J Bone Joint Surg Am* 2001;83:1656-9.

6

Spinal fusion for low-back pain: Systematic review on the accuracy of prognostic tests

Willems PC, Staal JB, Walenkamp GHIM, de Bie RA

Submitted

ABSTRACT

Background. Spinal fusion is a commonly used but controversial treatment for low-back pain, with clinical results that are comparable to those of programmed conservative care. As proper patient selection may improve these results, we investigated the prognostic accuracy of tests that clinicians use in daily practice to predict the outcome of spinal fusion for low-back pain.

Methods. We analysed studies that compared the results of prognostic tests (MRI, provocative discography, facet joint blocks, orthosis immobilisation, and temporary external fixation) with the clinical outcome of spinal fusion for low-back pain. Study populations with neurologic deficit, fracture, infection, ankylosing spondylitis, neoplasm or deformity were excluded. Ten studies met the inclusion criteria. Study quality was assessed by two independent reviewers with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist, modified for prognostic studies.

Findings. Three studies on immobilisation by an orthosis (median [range] positive likelihood ratio (LR), 1·10 [0·94 to 1·13], negative LR, 0·92 [0·39 to 1·12]), four studies on provocative discography (positive LR, 1·18 [0·70 to 1·71], negative LR, 0·74 [0·24 to 1·40]), and three studies on temporary external transpedicular fixation (positive LR, 1·22 [1·02 to 1·74], negative LR, 0·58 [0·15 to 0·94]) failed to show clinically useful prognostic accuracy. No studies reporting on MRI or facet joint blocks could meet the inclusion criteria. Obscure patient selection, high risk of verification bias, and outcome assessment with poorly validated instruments precluded strong conclusions. Results are applicable to secondary settings where patients with low-back pain are evaluated.

Interpretation. Results of spinal fusion for low-back pain remain unpredictable. Best evidence does not support the use of any prognostic test in clinical practice. At present, no subset of patients with low-back pain can be identified, for whom spinal fusion is a reliable and effective treatment.

INTRODUCTION

Although not life-threatening, low-back pain imposes huge costs to society,¹ either directly by healthcare consumption, or indirectly by lost productivity because of work absenteeism and early retirement². The estimated total costs of back pain in the USA have been reported to exceed 100 billion dollars per year³.

If conservative treatment programs fail,⁴ lumbar spinal fusion may be performed, however, its results are conflicting and hard to predict for the individual patient^{5,6}. Ten years ago lumbar fusion was not proven to be any better than placebo or natural history⁷. In 2001, the first randomised controlled trial (RCT)⁸ concluded that fusion does have a beneficial effect for some patients as compared to usual conservative care at two years follow-up, although the effect appeared to attenuate at 4 years follow-up⁹. Two later RCTs compared fusion to cognitive behavioural exercise therapy¹⁰ or an intensive rehabilitation program¹¹ and reported equal improvement for surgery and conservative treatment at one year and two years follow-up, respectively. As spinal fusion surgery is associated with substantial complications,¹² health care use,¹³ and morbidity,^{9,14} proper patient selection to identify those patients who will actually benefit from fusion, is mandatory.

History taking and physical examination may reveal psychological distress or social factors leading to chronicity of pain,¹⁵⁻¹⁷ however, they cannot predict the outcome of fusion¹⁸. The association of low-back pain with imaging is weak,^{19,20} and although MRI has been recommended as the imaging study of choice for the clinical evaluation of patients with back pain,²¹ its ability to predict the outcome of spinal fusion is unknown.

In practice, many clinicians rely on prognostic tests for patient selection. The most commonly used are provocative discography, facet joint blocks (both tests intend to identify the source of low-back pain), immobilisation by an orthosis, and temporary external transpedicular fixation of suspect spinal levels (both mechanical tests that intend to mimic the immobilising effect of a spinal fusion).

Considering that false-positive test results will lead to unnecessary invasive and expensive surgery with potential complications, whereas false-negative test results will withhold adequate treatment from patients who may benefit from fusion, it is essential to know the prognostic accuracy of these tests.

We therefore conducted a systematic review on the performance of MRI, provocative discography, facet joint blocks, orthosis immobilisation, and external fixation to predict the clinical outcome of lumbar spinal fusion for back pain. In addition, the risk of bias of the selected studies was appraised.

METHODS

Prognostic tests

The following tests were considered relevant for the purpose of this systematic review:

Magnetic Resonance Imaging (MRI)

Facet joint degeneration and abnormal disc morphology can be identified on MRI of the lumbar spine. Loss of T2-signal intensity, collapse, Modic changes²² and high intensity zones are commonly observed in the disc and presumed to be a source of pain²³.

Provocative discography

Under sterile conditions a stilette needle is advanced into the center of the intended disc space. Under fluoroscopic control a contrast agent is injected. If this injection provokes pain similar to the patient's usual pain, (i.e., concordant pain) and, preferably, if one or two control discs adjacent to the suspect disc do not elicit concordant pain, the test is considered positive²⁴. The extent of degeneration of the injected disc(s) is determined on fluoroscopy or a CT-scan immediately after the procedure.

Facet joint blocks

Using an aseptic technique and fluoroscopic guidance, local anesthetic is injected into the facet joint. Between 0.5 and 3 hours after injection, the amount of pain relief is recorded. In case of substantial pain relief, the test is considered positive²⁵.

Immobilisation by an orthosis

A standard brace or corset is prescribed, or a plaster cast can be applied²⁶. In a pantaloone cast, one hip is fixed within the cast for better immobilization of the lumbosacral junction. Patients are expected to wear the TLSO or cast for at least 2 to 4 weeks and are encouraged to perform as much daily life activities as possible. In case of significant pain relief, the test is considered positive.

Immobilisation by temporary external transpedicular fixation (TETF)

Under general anesthesia, antibiotic prophylaxis and fluoroscopy, two screws are

inserted percutaneously through the pedicles into the vertebra above, and two screws into the vertebra below the suspect disc(s), respectively. Postoperatively, the protruding screw ends are fixed externally with two vertical bars, which immobilises the disc(s) of interest. In case of adequate pain relief, the test is considered positive. Optionally, immobilisation can be discontinued without knowledge of the patient by fixing the bars horizontally (dynamisation),²⁷ which should annul pain relief.

Data Sources and Searches

A literature search was conducted according to the guidelines by Devillé et al²⁸. We explored Pubmed (1966-May 2010) and Embase (1974-May 2010) databases and used search terms for relevant test procedures, study design and patient population. For the tests the following terms were used: immobilisation, thoracolumbosacral orthosis, surgical cast(s), provocative discography, discography, temporary external fixation, facet joint blocks, zygapophyseal joint blocks, imaging and MRI. For study design we used the terms prognosis, prognostic, accuracy, predictive, diagnosis, diagnostic test(s), and diagnostic technique(s). For patient population the terms lumbar spine, lumbar vertebrae, lumbosacral, spinal, low back pain, degenerative disc disease, intervertebral disc(k), disc degeneration, failed back surgery syndrome, spondylosis, spinal fusion and spondylodesis. Both Mesh terms and free text words were entered.

Study selection

Two authors (PW, JBS) screened the titles and abstracts of all references identified by the literature search to determine whether they met the following inclusion criteria:

- Patients should suffer from low-back pain for at least three months.
- Studies should contain both patients with a positive and patients with a negative index test result, who subsequently underwent spinal fusion.
- Clinical outcome after fusion, which was considered as the reference standard, should be presented per individual patient in such a way that a relevant clinical improvement cut-off could be defined for analysis, and outcome could be dichotomized into "success" or "failure".
- Pain should have been incorporated as clinically-relevant outcome measure. Other possible outcome measures were subjective improvement, back-specific disability, disability for work, and patient satisfaction.

- Studies should include at least 20 patients.
- Only studies in English, German, French or Dutch were considered.
- Study populations with objective neurologic motor deficit, fracture, infectious disease, ankylosing spondylitis, neoplasm, congenital or adolescent idiopathic scoliosis, kyphosis, or adult scoliosis were excluded.

Full publications of studies which were considered as potentially relevant by both authors were retrieved. The articles were read and checked for final inclusion independently. Consensus meetings were organised to discuss any disagreement with regard to study selection. In cases where disagreement persisted, a third reviewer (RdB) was consulted for the final decision. The references of the articles identified by this search were checked for additional eligible studies.

Data Extraction and Quality Assessment

Relevant study data were retrieved by the same two reviewers using standardised forms. Extracted information included: standard reference data (first author, journal and publication year), number of patients, characteristics of study population before surgery (i.e., age, gender, severity and duration of pain and/or disability), index test of interest, spinal fusion method, measures of outcome, and clinical outcome.

The two reviewers independently assessed the methodological quality of included studies by means of a modified version of the Quality Assessment for Diagnostic Accuracy Studies (QUADAS) checklist²⁹. The QUADAS is a generally acknowledged checklist to assess the quality of primary studies of diagnostic accuracy. As there are no gold standard criteria for quality assessment of studies of prognostic accuracy, we modified the QUADAS checklist, as follows (see Table 1): Item 1 and 2 of the original QUADAS remained in the modified version. The original item 3 (*Is the reference standard likely to correctly classify the target condition?*) was left out because for the selected studies in the present review, the reference standard and target condition are the same (i.e., clinical outcome after fusion). Instead, whether the reference standard was assessed by valid measures of acceptable quality, was included as item 3. Item 4 of the original QUADAS (*Is the time period between the reference standard and the index test short enough to be reasonably sure that the target condition did not change in the time period between these tests?*) was removed because, in order to obtain a reliable estimation

Table 1. QUADAS modified for prognostic accuracy: Criteria list to assess the risk of bias of included studies

Quality criteria

1. Was the spectrum of patients representative of the patients who will receive the index test in practice?
2. Were selection criteria clearly described?
3. Were the outcomes used to assess recovery collected by means of validated measures of acceptable quality?*
4. Was a sufficiently long follow-up period (2 years or more) used to assess the outcome of the spinal fusion operation?*
5. Did all patients receive spinal fusion followed by the outcome assessment (i.e., the reference standard) regardless of the index test result?
6. Was the execution of the index test described in sufficient detail in order to permit replication of the test?
7. Was a clear cut-off point used to qualify positive versus negative results of the index test?*
8. Did the effect sizes that were used to consider patients as being recovered (i.e., the reference standard) meet accepted standards of clinical relevance, i.e., a minimal important change (MIC) of 30% or more?*
9. Were the clinical outcomes after fusion assessed without knowledge of the results of the index test?
10. Were the same clinical data available when index test results were interpreted as would be available when the test is used in practice?
11. Were uninterpretable results of the index test reported?
12. Were withdrawals from the study explained?

QUADAS = quality assessment of diagnostic accuracy studies⁵⁷

*Items 3, 4, 7, and 8 are modified items.

of clinical outcome after lumbar spinal fusion (i.e., reference standard), the length of follow-up should be at least two years³⁰ (modified item 4). The original item 5 (*Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?*) was left out because for inclusion in the present review, all analysed patients from the selected studies had undergone fusion and subsequent clinical outcome assessment. Item 6 of the original QUADAS (*Did patients receive the same reference standard regardless of the index test*

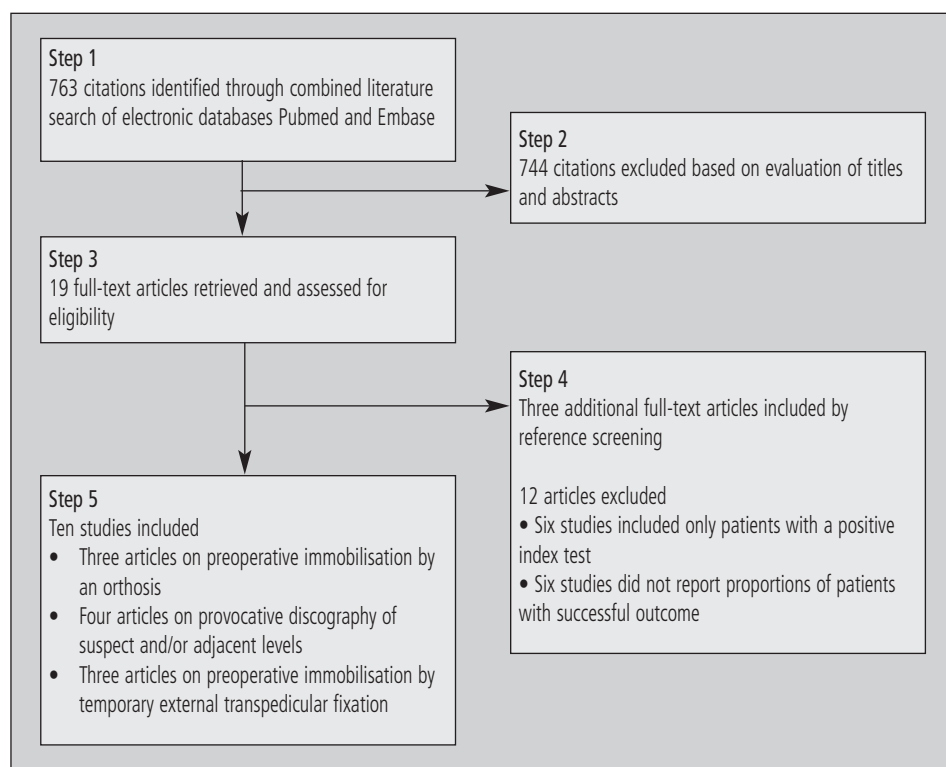
result?) remained unchanged (modified item 5). Item 7 of the original QUADAS (*Was the reference standard independent of the index test or did the index test form part of the reference standard?*) was removed because the outcome of fusion was assessed much later than the index test. The original item 8 was included in the modified QUADAS version as item 6. The modified item 7 was added to verify whether an objective, clearly defined cut-off point was mentioned to determine whether the index test was positive or negative. Item 9 of the original QUADAS (*Was the execution of the reference standard described in sufficient detail in order to permit replication of the test?*) was transformed into whether the assessment of clinical outcome after fusion was adequately addressed according to accepted standards of clinical importance³¹ (modified item 8). Item 10 of the original QUADAS (*Were the index test results interpreted without knowledge of the reference standard?*) was left out because fusion outcome was assessed much later than the preoperative index test. The original items 11, 12, 13 and 14 were included as items 9, 10, 11 and 12, respectively. We resolved disagreements about quality ratings by discussion and consensus. Studies that met more than half (seven or more) of the 12 criteria were classified “higher-quality”, and those that met six or less criteria “lower-quality”.

Data Synthesis and Analysis

By combining outcome (dichotomised into “success” or “failure”) with the test results (positive or negative), prognostic 2x2 tables with four cells (true positives, false negatives, false positives and true negatives) could be generated and test qualifiers, such as sensitivity, specificity, predictive values and likelihood ratios with 95% confidence intervals were calculated. Calculations were done with MetaDisc statistical software version 1.4³². Statistical pooling was only performed if studies on a specific index test were not hampered by statistical or clinical heterogeneity. Statistical heterogeneity was defined as non-overlapping 95% confidence intervals for estimates of sensitivity and specificity, and a difference in these estimates among the studies of more than 20%^{33,34}. We considered studies as clinically heterogeneous when patient groups, outcome measures, or the execution of index tests were different. In cases of statistical or clinical heterogeneity we refrained from statistical pooling and the results were presented per individual study.

RESULTS

Figure 1 shows the flow chart of studies from initial results of publication searches to final inclusion. Of the 22 selected full articles, we excluded six studies in which only patients with a positive index test had been selected for lumbar fusion³⁵⁻⁴⁰. We also excluded six other studies, in which prognostic accuracy could not be determined because only mean values of recovery were reported without proportions of patients with success or failure of treatment^{25,41-45}. Finally, 10 studies^{26,46-54} met the inclusion criteria.



Characteristics of included studies

Study characteristics are listed in Table 2. Three papers concerned immobilisation by an orthosis^{26,46,47}. Four papers reported on discography, of which two studies

Table 2. Characteristics of Primary Prognostic Studies on the Outcome of Lumbar Spinal Fusion

Source	Setting	N	Patients (N) for current analysis	Mean age ± sd (range, if specified)	Gender No. of male patients
Thoracolumbosacral orthosis					
Axelsson et al, ⁴⁷ 1995	Tertiary (University Hospital)	50	50	44 (20-68)	56 (28%)
Rask and Dall, ²⁷ 1993	Tertiary (University Hospital)	45	25	43.7 (20-61)	Not specified
Willems et al, ⁴⁸ 2006	Tertiary (Specialized Hospital)	257	107	40 ± 8.8 (range not specified)	39 (36%)
Provocative discography					
Colhoun et al, ⁵⁰ 1988	Tertiary (Orthopaedic Hospital)	195	168	39.1 (17-70)	86 (51%)
Esses et al, ⁵⁵ 1989	Tertiary (University Hospital)	32	22	41 (31-57)	18 (84%)

Systematic review on the accuracy of prognostic tests

Patient characteristics	Study design	Index test and criterion for positive result	Follow-up in months (range)	Method of fusion	Criteria for positive reference standard (=fusion outcome)
Intractable LBP: spondylolisthesis (n=24), facet/disc degeneration (n=11) or postlaminectomy syndrome (n=15), duration not specified Retrospective	Thoracolumbo-sacral orthosis or canvas corset; Positive in case of $\geq 50\%$ subjective pain relief	24	Postero-lateral fusion with autograft	Pain free or significantly improved on a 5-point pain scale and satisfied on a 3-point satisfaction scale*	
> 6 months back pain (mean 3.9 years), no neurological motor deficit, herniation or olisthesis >2mm, 38% with prior spine surgery	Retrospective	Fiberglass pantaloons cast; Positive in case of pain relief that returned after removal of the cast	Minimum of 6	Postero-lateral fusion with autograft	Significant subjective pain relief*
Incapacitating LBP, mean 3.7 (0.5-31) years, no neurological motor deficit and routine testing indecisive, 65% with prior spine surgery (N=70)	Retrospective	Pantaloons plaster cast; Positive in case of subjective substantial pain relief in the cast	Median of 76 (15-144)	Postero-lateral (n=79) or anterior fusion (n=28)	$\geq 30\%$ decrease in pain on a VAS (0-100)
Persistent LBP, no previous back surgery, duration not specified	Retrospective	Provocative discography; Positive in case of typical pain reproduction which was not present in adjacent control discs	Mean of 43 (24-120)	Anterior or posterior fusion (numbers not specified)	Complete pain relief or significant subjective improvement, resumption of work/normal duties and no intake of analgesics*
Longstanding LBP, mean duration of 6.2 (1-20) years, 54% with prior spine surgery	Prospective	Provocative discography, no control discs; Positive in case of typical pain reproduction	"When fusion was noted"	Postero-lateral fusion with autograft	Complete or significant relief of pain*

Chapter 6

Table 2. Characteristics of Primary Prognostic Studies on the Outcome of Lumbar Spinal Fusion

Source	Setting	N	Patients (N) for current analysis	Mean age \pm sd (range, if specified)	Gender No. of male patients
Gill and Blumenthal, ⁵³ 1991	Tertiary (Orthopaedic Institute)	53	53	34 (21-50)	36 (68%)
Willems et al, ⁴⁹ 2007	Tertiary (Specialized Hospital)	197	82	40 \pm 8.5 (range not specified)	26 (32%)
Temporary external transpedicular fixation					
Elmans et al, ⁵¹ 2005	Tertiary (Specialized Hospital)	330	123	42 \pm 9 (range not specified)	45 (37%)
Heini et al, ⁵² 2004	Tertiary (University Hospital)	63	36	48 (26-67)	22 (62%)
Jeanneret et al, ⁵⁴ 1994	Secondary (Regional Hospital)	101	43	48 (22-74)	43 (58%)

Systematic review on the accuracy of prognostic tests

Patient characteristics	Study design	Index test and criterion for positive result	Follow-up in months (range)	Method of fusion	Criteria for positive reference standard (=fusion outcome)
LBP with a mean disability of 11 (3-120) months, all selected by concordant pain response provoked at discography L5-S1	Design not specified	Discography image of L5-S1, no control disc(s); Positive in case of annular tear extending to the periphery	Mean of 36 (24-54)	Anterior fusion with allograft (n=48) or autograft (n=5)	Relief of $\geq 75\%$ of initial back pain on VAS, return to work and no use of narcotics
Incapacitating LBP > 1 year (mean duration not specified), no neurological motor deficit and routine testing indecisive, 65% with prior spine surgery (N=53)	Retrospective	Provocative discography of adjacent to intended fusion; positive in case of no or unfamiliar pain reproduction	Mean of 80 (15-144)	Postero-lateral or anterior fusion (numbers not specified)	$\geq 30\%$ decrease in pain on a VAS (0-100)
Incapacitating LBP > 1 year (mean duration of 6 ± 5 years) with inconclusive routine testing, 62% with prior spine surgery	Prospective	TETF with dynamisation; Positive if VAS in placebo was ≥ 30 points more than VAS in fixation	Median of 79 (15-44) months	Anterior (n=33) or postero-lateral (n=90) fusion	$\geq 30\%$ decrease in pain on a VAS (0-100)
LBP with a mean duration of 5 (1-20) years, 67% with prior spine surgery	Prospective	TETF without dynamisation; Positive if pain on VAS and use of analgesics decreased sufficiently (estimated by surgeon)	Mean of 32 (23-60)	Postero-lateral fusion, except 3 dynamic fixations and 1 anterior fusion	No or little pain on a VAS and no pain medication*
Chronic LBP with or without leg pain, duration not specified, routine testing indeterminate	Design not specified	TETF with dynamisation; Positive if pain was reduced at stabilization and returned at destabilization	Mean of 50 (24-92)	Posterior or anterior fusion (numbers not specified)	Almost completely pain free, no pain medication and working*

focused on provocative discography of suspect levels,^{49,54} one study on provocation of the levels adjacent to the intended fusion,⁴⁸ and the fourth study focused on the amount of degeneration as registered at discography in a group of patients with a positive discographic pain response⁵². Three papers evaluated trial-immobilisation by external fixation^{50,51,53}.

The sample sizes of included studies ranged from 22 to 162. Two studies reported exclusively on patients without prior spine surgery^{49,52}. The length of follow-up ranged from six months or "when fusion was noted"⁵⁴ to 12 years. Either anterior or posterolateral fusion was performed. In those studies in which both procedures were performed, no difference in outcome between the two types of fusion was reported⁴⁷⁻⁵⁰.

Five studies^{47,48,50-52} used a VAS to score pain: Three studies^{47,48,50} defined a cut-off point of at least 30% decrease in pain as a clinically relevant improvement, one study⁵² a decrease of at least 75%, and one study⁵¹ "little or no pain on a

Table 3. Risk of bias: Number of quality criteria of the QUADAS Checklist modified for prognostic accuracy, that were met by each included study (+ = yes; - = no; ? = unclear)

Quality Criteria						
Source	1	2	3	4	5	6
Thoracolumbosacral orthosis						
Axelsson, ⁴⁷ 1995	+	-	-	+	+	+
Rask and Dall, ²⁷ 1993	+	+	-	-	-	+
Willems, ⁴⁸ 2006	-	+	+	+	-	+
Provocative discography						
Colhoun, ⁵⁰ 1988	+	?	-	+	+	+
Esses, ⁵⁵ 1989	+	-	-	-	-	-
Gill and Blumenthal, ⁵³ 1991	-	-	?	+	+	?
Willems, ⁴⁹ 2007	-	+	+	+	-	+
Temporary external transpedicular fixation						
Elmans, ⁵¹ 2005	-	-	+	+	-	+
Heini, ⁵² 2004	-	-	+	+	-	+
Jeanneret, ⁵⁴ 1994	+	?	-	+	-	+
Number of trials meeting criteria	5	3	4	8	3	8
*Maximum score is 12						

LBP = low-back pain; TETF = temporary external transpedicular fixation; VAS = visual analog scale

*No validated outcome measure reported

VAS". The other studies used a subjective pain scale ("pain free" or "significant pain relief" versus "insignificant or no pain relief")^{26,46,49,53,54}.

No studies reporting on facet joint blocks or MRI could satisfy our inclusion criteria.

Methodological Quality

The two reviewers agreed on 77 of the 120 items (64%) scored. After discussion, consensus could be reached on all items. Most disagreements were because of reading errors or ambiguous reporting. The results of the quality assessment are shown in Table 3. Three studies received a positive assessment of meeting seven or more modified QUADAS criteria, and were classified as "higher-quality"^{47,48,50}. The most prevalent shortcomings were: "Uninterpretable index test results not reported", "no clear cut-off point defined for positive and negative index test results", and "not all tested patients underwent fusion regardless of the index test result (verification bias)".

7	8	9	10	11	12	Total score*
+	-	?	+	?	?	6
-	-	+	+	-	+	6
-	+	+	+	-	+	8
-	-	?	+	?	?	5
?	-	-	+	-	?	2
+	-	-	?	?	+	4
?	+	+	-	-	+	7
+	+	+	?	?	+	7
-	-	-	+	+	+	6
-	-	-	+	-	?	4
3	3	4	7	1	6	

Prognostic accuracy

Statistical pooling was not feasible because of heterogenous patient populations, different test protocols, poor reporting, and variability in outcome measures.

Orthosis. An orthosis or cast immobilization could neither confirm nor rule out a good outcome after spinal fusion with variable sensitivity (0.43-0.94; median [range] positive LR, 1.10 [0.94-1.13]) and specificity (0.14-0.61; median [range] negative LR, 0.92 [0.39-1.12]), respectively^{26,46,47}.

Provocative discography. The four studies reporting on the prognostic value of discography showed variable sensitivity (range 0.40-0.88; median [range] positive likelihood ratio (LR), 1.18 [0.70-1.71]) and specificity (0.27-0.48; median [range] negative LR, 0.74 [0.24-1.40])^{48,49,52,54} (Table 4). In one study likelihood ratios were

Table 4. Summary Accuracy of orthosis, provocative discography, and temporary external transpedicular fixation for predicting the outcome of lumbar spinal fusion

Source	Sample size	Sensitivity	Specificity	Positive predictive	Negative predictive value	Positive LR (95% CI) value	Negative LR (95% CI)
Thoracolumbosacral orthosis							
Axelsson, ⁴⁷ 1995	50	0.61	0.35	0.64	0.32	0.94 (0.60 to 1.5)	1.1 (0.52 to 2.4)
Rask and Dall, ²⁷ 1993	25	0.94	0.14	0.74	0.50	1.1 (0.80 to 1.5)	0.39 (0.03 to 5.4)
Willems, ⁴⁸ 2006	107	0.43	0.61	0.44	0.61	1.1 (0.71 to 1.8)	0.92 (0.67 to 1.3)
Provocative discography							
Colhoun, ⁵⁰ 1988	168	0.88	0.48	0.88	0.48	1.7 (1.2 to 2.4)	0.24 (0.13 to 0.43)
Esses, ⁵⁵ 1989	22	0.40	0.43	0.60	0.25	0.70 (0.29 to 1.7)	1.4 (0.54 to 3.6)
Gill and Blumenthal, ⁵³ 1991	53	0.81	0.41	0.74	0.50	1.4 (0.89 to 2.1)	0.47 (0.20 to 1.1)
Willems, ⁴⁹ 2007	82	0.73	0.27	0.45	0.55	0.99 (0.76 to 1.3)	1.0 (0.52 to 1.9)
Temporary external transpedicular fixation							
Elmans, ⁵¹ 2005	123	0.80	0.34	0.46	0.71	1.22 (0.98 to 1.5)	0.58 (0.31 to 1.1)
Heini, ⁵² 2004	36	0.81	0.20	0.45	0.57	1.0 (0.74 to 1.4)	0.94 (0.24 to 3.6)
Jeanneret, ⁵⁴ 1994	43	0.93	0.47	0.77	0.78	1.7 (1.1 to 2.8)	0.15 (0.03 to 0.65)

CI = confidence interval; LR = likelihood ratio.

statistically significant⁴⁹ (positive LR, 1.71; 95% CI, 1.21-2.4 and negative LR, 0.24; 95% CI, 0.13-0.43), but specificity was low, 0.48, meaning that only half of the patients who would not improve after fusion, could be detected.

Temporary external transpedicular fixation. Sensitivity was generally high (range, 0.80-0.93; median [range] positive LR, 1.22 [1.02-1.74]), but specificity was low (range 0.20-0.47; median [range] negative LR, 0.58 [0.15-0.94])^{50,51,53, 54} In one study likelihood ratios were statistically significant⁵³ (positive LR, 1.74; 95% CI, 1.07-2.8 and negative LR, 0.15; 95% CI, 0.03-0.65) with a specificity of 0.47.

DISCUSSION

We systematically reviewed the reported accuracy of prognostic tests that are commonly used in clinical practice to predict the outcome of spinal fusion for low-back pain. From the vast amount of literature on the management of back pain, only ten studies evaluating three tests, allowed determination of prognostic accuracy. All tests failed to reliably predict the outcome of spinal fusion with likelihood ratios approaching one (no discriminative value at all).

The relatively small number of eligible studies and the substantial risks of bias in most selected studies, limited our findings. A major drawback was that in all but three studies, a proportion of patients who had undergone the index test with a negative test result, had been denied fusion and excluded from analysis (risk of verification bias). Ideally, fusion should have been performed independent of the index test result.

Our analysis was limited to pain as this was the only measure of outcome to be consistently incorporated in all studies. Although pain represents only one aspect of the complex low-back pain problem, it is the primary indication for operative treatment.

In order to determine the accuracy of a test, it is essential that the interpretation of the test result is unequivocal and that a clear cut-off point for a positive versus negative test result has been defined. The lack of a clear cut-off point in most studies has hampered pooling of results and impairs replication of the test in other settings.

Six studies^{25,41-45} that reported mean clinical improvement after fusion for groups of patients with a positive and a negative index test result, respectively, were excluded from our review. Since no proportions of patients with a good clinical outcome were reported,³¹ no 2x2 tables could be created to determine prognostic accuracy. Three of these excluded studies reported on MRI with conflicting results: In one study⁴¹

inflammatory vertebral endplate changes (Modic type I)²² were significantly related to continued back pain after fusion, whereas the other two studies showed a significantly better,⁴³ or a “relatively” better (no statistics)⁴² outcome for patients with Modic type I endplate changes, respectively. A study on preoperative test bracing,⁴⁴ revealed “a clear tendency for poorer prognosis for patients who had responded poorly to the brace”. Another excluded study focusing on pressure-controlled discography,⁴⁵ reported no significant differences in long-term surgical outcome across the entire sample. A study on facet joint blocks²⁵ failed to show any significant correlation between test results and the outcome of spinal fusion.

The lack of scientifically proven accuracy of prognostic tests is reflected in the high degree of clinical uncertainty and variability in outcome of fusion surgery for low-back pain^{5,55}. Studies among spine surgeons show that there is no consensus in treatment strategy,^{56,57} and the present results confirm that in many clinical practices patients are scheduled for fusion on the basis of tests, of which the accuracy is insufficient or at best unknown. In this respect, it is disappointing that for MRI, the recommended standard diagnostic tool for low-back pain,⁵⁸ no studies could be identified to determine its prognostic accuracy for the outcome of spinal fusion.

As treatment for low-back pain, cognitive behavioural therapy or intensive exercise programs^{10,11,59} have similar results but with considerably less complications, morbidity and costs,^{9,14} as compared to fusion. As patients with a good outcome cannot be identified, fusion should not be recommended as standard treatment for low-back pain. However, if we feel spinal fusion may be effective for some patients with persisting symptoms after programmed conservative care, future prognostic studies with larger patient numbers, a clear description of participants, blinding of assessment and independent interpretation of test results are needed. These studies should compare both positively and negatively tested patients for whom clinical outcome is defined by a consensus cut-off point of improvement in pain and functional status, a so-called minimal clinically important change^{31,60}. Thus, reliable predictors of intervention would enable physicians to counsel their patients better in weighing the risks and benefits of fusion. For now, however, best evidence does not support the use of any prognostic test in clinical practice. No subset of patients with low-back pain could be identified for whom spinal fusion is a reliable and effective treatment.

REFERENCES

1. Maniadakis N, Gray A. The economic burden of back pain in the UK. *Pain*. Jan 2000;84(1):95-103.
2. van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. *Pain*. Aug 1995;62(2):233-240.
3. Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *J Bone Joint Surg Am*. Apr 2006;88 Suppl 2:21-24.
4. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*. Mar 2006;15 Suppl 2:S192-300.
5. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery - the case for restraint. *N Engl J Med*. Feb 12 2004;350(7):722-726.
6. Krismer M. Fusion of the lumbar spine. A consideration of the indications. *J Bone Joint Surg Br*. Aug 2002;84(6):783-794.
7. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine*. Sep 1 1999;24(17):1820-1832.
8. Fritzell P, Hagg O, Wessberg P, Nordwall A. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine*. Dec 1 2001;26(23):2521-2532; discussion 2532-2524.
9. Fritzell P, Hagg O, Jonsson D, Nordwall A. Cost-effectiveness of lumbar fusion and nonsurgical treatment for chronic low back pain in the Swedish Lumbar Spine Study: a multicenter, randomized, controlled trial from the Swedish Lumbar Spine Study Group. *Spine*. Feb 15 2004;29(4):421-434; discussion Z423.
10. Brox JI, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine*. Sep 1 2003;28(17):1913-1921.
11. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *Bmj*. May 28 2005;330(7502):1233.

12. Wilson-MacDonald J, Fairbank J, Frost H, et al. The MRC spine stabilization trial: surgical methods, outcomes, costs, and complications of surgical stabilization. *Spine*. Oct 1 2008;33(21):2334-2340.
13. Rivero-Arias O, Campbell H, Gray A, Fairbank J, Frost H, Wilson-MacDonald J. Surgical stabilisation of the spine compared with a programme of intensive rehabilitation for the management of patients with chronic low back pain: cost utility analysis based on a randomised controlled trial. *Bmj*. May 28 2005;330(7502):1239.
14. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. *Jama*. Apr 7;303(13):1259-1265.
15. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *Jama*. Apr 7;303(13):1295-1302.
16. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *Jama*. Aug 12 1992;268(6):760-765.
17. Nelson MA, Allen P, Clamp SE, de Dombal FT. Reliability and reproducibility of clinical findings in low-back pain. *Spine (Phila Pa 1976)*. Mar-Apr 1979;4(2):97-101.
18. Carragee EJ. Clinical practice. Persistent low back pain. *N Engl J Med*. May 5 2005;352(18):1891-1898.
19. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine*. Feb 15 1997;22(4):427-434.
20. Kumar MN, Jacquot F, Hall H. Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease. *Eur Spine J*. Aug 2001;10(4):309-313.
21. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 6: magnetic resonance imaging and discography for patient selection for lumbar fusion. *J Neurosurg Spine*. Jun 2005;2(6):662-669.
22. Modic MT, Masaryk TJ, Ross JS, Carter JR. Imaging of degenerative disk disease. *Radiology*. Jul 1988;168(1):177-186.
23. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila Pa 1976)*. Sep 1 2001;26(17):1873-1878.

24. Guyer RD, Ohnmeiss DD. Lumbar discography. Position statement from the North American Spine Society Diagnostic and Therapeutic Committee. *Spine (Phila Pa 1976)*. Sep 15 1995;20(18):2048-2059.
25. Esses SI, Moro JK. The value of facet joint blocks in patient selection for lumbar fusion. *Spine*. Feb 1993;18(2):185-190.
26. Rask B, Dall BE. Use of the pantaloons cast for the selection of fusion candidates in the treatment of chronic low back pain. *Clin Orthop Relat Res*. Mar 1993(288):148-157.
27. van der Schaaf DB, van Limbeek J, Pavlov PW. Temporary external transpedicular fixation of the lumbosacral spine. *Spine (Phila Pa 1976)*. Mar 1 1999;24(5):481-484; discussion 484-485.
28. Deville WL, Buntinx F, Bouter LM, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol*. Jul 3 2002;2:9.
29. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. Nov 10 2003;3:25.
30. Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. *Jama*. May 25 1994;271(20):1609-1614.
31. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*. Jan 1 2008;33(1):90-94.
32. Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31.
33. Jellema P, van der Windt DA, Bruinvels DJ, et al. Value of symptoms and additional diagnostic tests for colorectal cancer in primary care: systematic review and meta-analysis. *Bmj*.340:c1269.
34. van der Windt DA, Jellema P, Mulder CJ, Kneepkens CM, van der Horst HE. Diagnostic testing for celiac disease among patients with abdominal symptoms: a systematic review. *Jama*. May 5;303(17):1738-1746.
35. Ohtori S, Kinoshita T, Yamashita M, et al. Results of surgery for discogenic low back pain: a randomized study using discography versus discoblock for diagnosis. *Spine (Phila Pa 1976)*. Jun 1 2009;34(13):1345-1348.
36. Lovely TJ, Rastogi P. The value of provocative facet blocking as a predictor of success in lumbar spine fusion. *J Spinal Disord*. Dec 1997;10(6):512-517.

37. Bednar DA, Raducan V. External spinal skeletal fixation in the management of back pain. *Clin Orthop Relat Res*. Jan 1996(322):131-139.
38. Axelsson P, Johnsson R, Stromqvist B, Andreasson H. Temporary external pedicular fixation versus definitive bony fusion: a prospective comparative study on pain relief and function. *Eur Spine J*. Feb 2003;12(1):41-47.
39. Madan S, Gundanna M, Harley JM, Boeree NR, Sampson M. Does provocative discography screening of discogenic back pain improve surgical outcome? *J Spinal Disord Tech*. Jun 2002;15(3):245-251.
40. Peng B, Chen J, Kuang Z, Li D, Pang X, Zhang X. Diagnosis and surgical treatment of back pain originating from endplate. *Eur Spine J*. Jul 2009;18(7):1035-1040.
41. Buttermann GR, Heithoff KB, Ogilvie JW, Transfeldt EE, Cohen M. Vertebral body MRI related to lumbar fusion results. *Eur Spine J*. 1997;6(2):115-120.
42. Chataigner H, Onimus M, Polette A. [Surgery for degenerative lumbar disc disease. Should the black disc be grafted?]. *Rev Chir Orthop Reparatrice Appar Mot*. Nov 1998;84(7):583-589.
43. Esposito P, Pinheiro-Franco JL, Froelich S, Maitrot D. Predictive value of MRI vertebral endplate signal changes (Modic) on outcome of surgically treated degenerative disc disease. Results of a cohort study including 60 patients. *Neurochirurgie*. Sep 2006;52(4):315-322.
44. Christensen FB, Karlsmose B, Hansen ES, Bunger CE. Radiological and functional outcome after anterior lumbar interbody spinal fusion. *Eur Spine J*. 1996;5(5):293-298.
45. Derby R, Howard MW, Grant JM, Lettice JJ, Van Peteghem PK, Ryan DP. The ability of pressure-controlled discography to predict surgical and nonsurgical outcomes. *Spine (Phila Pa 1976)*. Feb 15 1999;24(4):364-371; discussion 371-362.
46. Axelsson P, Johnsson R, Stromqvist B, Nilsson LT, Akesson M. Orthosis as prognostic instrument in lumbar fusion: no predictive value in 50 cases followed prospectively. *J Spinal Disord*. Aug 1995;8(4):284-288.
47. Willems PC, Elmans L, Anderson PG, Jacobs WC, van der Schaaf DB, de Kleuver M. The value of a pantaloon cast test in surgical decision making for chronic low back pain patients: a systematic review of the literature supplemented with a prospective cohort study. *Eur Spine J*. Oct 2006;15(10):1487-1494.
48. Willems PC, Elmans L, Anderson PG, van der Schaaf DB, de Kleuver M. Provocative discography and lumbar fusion: is preoperative assessment of adjacent discs useful? *Spine (Phila Pa 1976)*. May 1 2007;32(10):1094-1099; discussion 1100.

49. Colhoun E, McCall IW, Williams L, Cassar Pullicino VN. Provocation discography as a guide to planning operations on the spine. *J Bone Joint Surg Br.* Mar 1988;70(2):267-271.
50. Elmans L, Willems PC, Anderson PG, van Limbeek J, de Kleuver M, van der Schaaf D. Temporary external transpedicular fixation of the lumbosacral spine: a prospective, longitudinal study in 330 patients. *Spine.* Dec 15 2005;30(24):2813-2816.
51. Heini PF, Gahrlich U, Orler R. The external fixator: a tool for evaluation of complex low back pain problems. *J Spinal Disord Tech.* Feb 2004;17(1):8-14.
52. Gill K, Blumenthal SL. Functional results after anterior lumbar fusion at L5-S1 in patients with normal and abnormal MRI scans. *Spine.* Aug 1992;17(8):940-942.
53. Jeanneret B, Jovanovic M, Magerl F. Percutaneous diagnostic stabilization for low back pain. Correlation with results after fusion operations. *Clin Orthop Relat Res.* Jul 1994(304):130-138.
54. Esses SI, Botsford DJ, Kostuik JP. The role of external spinal skeletal fixation in the assessment of low-back disorders. *Spine (Phila Pa 1976).* Jun 1989;14(6):594-601.
55. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES. United States' trends and regional variations in lumbar spine surgery: 1992-2003. *Spine.* Nov 1 2006;31(23):2707-2714.
56. Irwin ZN, Hilibrand A, Gustavel M, et al. Variation in surgical decision making for degenerative spinal disorders. Part I: lumbar spine. *Spine (Phila Pa 1976).* Oct 1 2005;30(19):2208-2213.
57. Katz JN, Lipson SJ, Lew RA, et al. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. *Spine (Phila Pa 1976).* May 15 1997;22(10):1123-1131.
58. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 2: assessment of functional outcome. *J Neurosurg Spine.* Jun 2005;2(6):639-646.
59. Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary rehabilitation for chronic low back pain: systematic review. *Bmj.* Jun 23 2001;322(7301):1511-1516.
60. van der Roer N, Ostelo RW, Bekkering GE, van Tulder MW, de Vet HC. Minimal clinically important change for pain intensity, functional status, and general health status in patients with nonspecific low back pain. *Spine (Phila Pa 1976).* Mar 1 2006;31(5):578-582.

GENERAL DISCUSSION

Despite greater knowledge, expertise, and health care resources for treating spinal pathologies, chronic disability resulting from nonspecific low back pain has risen exponentially in western society. Surgical intervention by means of lumbar spinal fusion remains a controversial treatment with variable and unpredictable results for the individual patient¹². Whereas ten years ago there was no evidence that lumbar fusion might be any better than placebo or natural history¹⁷, the first randomized controlled trial (RCT)¹⁶, which was performed in 2001, indicated that fusion does have a beneficial effect for some patients as compared to usual conservative care, although the effect appeared to attenuate at longer follow-up¹⁵. Two more recent RCTs compared fusion surgery to cognitive behavioural based exercise therapy⁶ or an intensive rehabilitation program¹³. Both studies reported equal improvement for the surgical and the conservative treatment modality at 1 year and 2 year follow-up, respectively. Fusion, however, had a higher complication rate³⁸ and appeared to be less cost- effective than intensive rehabilitation²⁶.

When aiming to improve the results of fusion for chronic low back pain (CLBP), it is crucial to know which patients will benefit from fusion. Unfortunately, a proper set of tools for patient selection does not exist. History taking and physical examination are helpful in identifying signs of sciatica or serious spinal conditions, such as fracture, tumor, infection or deformity. Psychosocial risk factors for delayed recovery may be identified as well, however, there are no specific findings to predict the outcome of lumbar fusion in CLBP patients⁸. Imaging generally reveals nonspecific findings of degeneration, such as annular disruption^{3,9,21} and endplate changes³⁷, which have been associated with low back pain, but are also commonly found in cross-sectional studies of asymptomatic persons^{5,22}. Essentially, in the evaluation of low back pain, the indications for imaging are not to confirm nonspecific LBP or to predict the outcome of its treatment, but to rule out specific causes, such as fractures, infection, tumors or impending neurological injury.

Aims: Prognostic tests for spinal fusion in CLBP

The aims of the present thesis were to investigate the value, use and appreciation of three prognostic tests (i.e., immobilization by a thoracolumbosacral orthosis (TLSO), provocative discography, and trial immobilization by temporary external transpedicular fixation (TETF)), which are commonly used in daily practice for patient selection. These tests aim to predict the outcome of lumbar spinal fusion either by mimicking a fusion, or by trying to identify the source of pain in the individual patient.

Aim 1: Is there consensus among spine surgeons regarding the use of prognostic tests for lumbar spinal fusion in CLBP patients?

In chapter 2, the results of a nationwide survey among spine surgeons in the Netherlands were presented. The comments were compared with findings from the prevailing literature. The survey revealed a lack of uniformity in the use and appreciation of prognostic tests for lumbar fusion in CLBP patients. Prognostic factors known from the literature were not consistently incorporated in the surgeons' decision making process either.

Aim 2: Does immobilization by a pantaloone cast truly minimize lumbosacral motion?

In chapter 3a, the hypothesized working mechanism of a pantaloone cast (minimization of lumbosacral joint mobility) was studied. In patients who were admitted for a temporary external transpedicular fixation test (TETF), infrared light markers were fixed to the protruding pins of transpedicular screws fixed at two spinal levels. Three-dimensional motion between these levels was analyzed opto-electronically. During dynamic test conditions (walking), wearing a pantaloone cast did not significantly decrease lumbosacral joint motion.

Aim 3: Can immobilization by a thoracolumbosacral orthosis (TLSO) predict the clinical outcome of spinal fusion for CLBP?

In chapter 3b, a systematic review of the literature supplemented with a prospective cohort study was performed in order to assess the value of a TLSO in surgical decision-making. It appeared that only in a small group of CLBP patients with no

prior spine surgery, a pantaloone cast test with substantial pain relief suggests a favorable outcome of lumbar fusion as compared to conservative treatment. In patients with prior spine surgery the test was of no value.

Aim 4: Does provocative discography of adjacent segments predict the long-term clinical outcome of spinal fusion for CLBP?

In chapter 4a, a cohort of patients for whom the decision to perform lumbar fusion was based on a TETF trial, was studied. The results of preoperative discography of levels adjacent to the fusion were analyzed retrospectively. It was concluded that in this select group of patients the discographic status of discs adjacent to a lumbar fusion could not predict the clinical outcome of fusion at long-term follow-up.

Aim 5: What is the incidence of postdiscography discitis and is there a need for routine antibiotic prophylaxis?

In search for clinical guidelines, the risk of postdiscography discitis was studied in chapter 4b, by means of a systematic literature review and a cohort of 200 consecutive patients. Without the use of prophylactic antibiotics, an overall incidence of discitis of 0.25% was found in the literature. In the patient cohort no discitis was noted. To prove that antibiotics actually prevent postdiscography discitis, a randomized trial of 9,000 patients would be needed to reach significance. Given the possible adverse effects of antibiotics, it was concluded that the routine use of prophylactic antibiotics in lumbar discography is not indicated.

Aim 6: Can temporary external transpedicular fixation (TETF) help to predict the clinical outcome of spinal fusion in CLBP patients with an equivocal indication for surgery?

In chapter 5, TETF as a test to predict the long-term clinical outcome of lumbar fusion was studied in a group of patients for whom there was doubt about the indication for surgery. Using strict and objective criteria of pain reduction on a visual analogue scale, the TETF test failed to predict the clinical outcome of fusion in this select group of patients. It was concluded that in chronic low back pain patients with a doubtful indication for fusion, TETF is not to be recommended as a supplemental tool for surgical decision-making.

Aim 7: What is the prognostic accuracy of the most commonly used tests in clinical practice to predict the outcome of spinal fusion for CLBP?

Chapter 6 presents the results of a systematic literature review concerning the prognostic accuracy of magnetic resonance imaging (MRI), TLSO immobilization, TETF, provocative discography and facet joint infiltration. Only 10 studies reporting on three different index tests (discography, TLSO immobilization and TETF) could be selected. It appeared that the overall methodological quality of the included studies was poor and the determined accuracy of the prognostic tests was low with likelihood ratios approaching 1 in most studies. It was concluded that best evidence does not support the value of any of these tests to improve patient selection in daily practice.

Based on the studies in the present thesis, there are no reliable prognostic tests at present, to aid surgeons in the clinical decision whether or not to perform a lumbar fusion in a patient with CLBP.

If it appears from the literature that lumbar spinal fusion as a treatment for CLBP is no better or worse than structured conservative care programs, and if current prognostic tests fail to improve overall results by better patient selection, should spinal fusion still be considered as a standard treatment for CLBP?

It has been suggested that the elimination of segmental motion by spinal fusion creates an abnormal biomechanical condition leading to degeneration of adjacent segments and subsequent re-operations⁴. In that respect, motion preserving strategies such as Total Disc Replacement (TDR) should, theoretically, provide better long-term clinical outcome. The use of TDR, however, appears to be limited to a relatively small group of patients with early disc degeneration with preservation of disc height and in the absence of significant facet joint arthritis¹⁴. In these highly selected patients the clinical results of TDR at 2 and at 5 years follow-up, respectively, have not been proven to be superior to lumbar fusion¹⁸. Moreover, the long-term endurance of TDR and the fate of the facet joints following these implants are still unknown. The rate of technically difficult revision procedures with serious complications is likely to increase at longer-term follow-up^{10,33}. In conclusion, at this stage TDR cannot be considered a more appropriate treatment for CLBP as compared to spinal fusion.

Approach of CLBP patients in daily practice

What could be the reason that the results of lumbar fusion and Total Disc Replacement for chronic low back pain remain disappointing despite tremendous efforts in research and innovation?

Basically, the surgical point of view, i.e., the search for an underlying anatomic cause or pain source, should not be our first approach in the evaluation and treatment of CLBP. The traditional biomedical treatment model does not appear to work well in these patients. This model suggests the recognition of symptoms and signs, identification of underlying injury or disease, treatment with specific therapy and the expectation that the patient will recover. Such an approach indeed works well when dealing with patients with specific LBP in whom causality for the pain can be identified, such as in spondylolisthesis, tumor or trauma. In the case of CLBP, however, causality is more complex and back pain is rather a symptom than a diagnosis in patients with different stages of impairment and disability. These patients appear to be better evaluated by the so-called biopsychosocial model³⁶, which posits that biologic, social and psychological factors all are implicated in any given state of health or disease. The model aims to identify factors associated with delayed recovery³⁵ and calls for assessment of cultural beliefs, anxiety, kinesiphobia, depression, stress from work or family, job dissatisfaction, somatisation, lack of control, and a high perception of disability. All these factors may affect the patient's symptoms and resulting disability. Aiming to treat underlying psychosocial factors and empowering the patient to take responsibility for managing his or her own condition, should be key factors in the initial treatment of CLBP patients. Chronic pain has much less impact on the well-being of a patient if he or she can function normally. Treatment should thus primarily focus on functional capacity with return to work and usual activities, and secondarily on pain relief.

In recent years, much progress has been made with cognitive behavioral and intensive exercise programs. The goal of cognitive behavioral interventions is to have the patient understand, accept, and gain control of his or her back pain problem and possible consequences, such as loss of self-esteem, fear of movement, depression, family problems, work loss and social withdrawal. Behavioral therapy has been proven to be more effective than placebo or no treatment² and there is

strong evidence that it should be used in patients with CLBP²⁴, especially if psychosocial signs or symptoms are present. In occupational settings, it has been shown that individual graded behavioral-oriented exercise programs can significantly reduce days lost at work^{30,32}.

Different kinds of exercise therapies appear to be equally effective for CLBP^{1,34}. However, it has been shown that intensive multidisciplinary biopsychosocial rehabilitation with a functional restoration approach reduces pain significantly more than inpatient or outpatient non-multidisciplinary programs or usual care^{8,19}.

Future care for CLBP patients: The case for a multidisciplinary Spine Center

In current medical practice, patients seeking care can contact clinicians from different medical disciplines with expertise at treating CLBP. These doctors all have their own specific training skills, opinions and therapy proposals, which may differ substantially. There is not even assurance that clinicians within the same medical discipline will offer a similar treatment. This can be both confusing and frustrating for patients. Ideally, clinicians should be aware of all commonly used therapies and able to counsel their patients which treatment may be preferable at that moment. Consequently, they should be willing to refer patients to colleagues, either within or outside their own discipline, who may be better trained to provide the most appropriate treatment²⁰.

Current evidence points to improved triage of low back pain patients in a multidisciplinary setting or Spine Center with a standardized intake, which includes thorough history taking, physical examination, psychosocial evaluation and adequate imaging. Patients expect diagnostic explanations, adequate pain relief from treatment, instructions on how to deal with persisting pain and disability, and perhaps help with sickness certification. Good patient counselling that provides a reliable estimation of the effectiveness of a planned therapy, should be evidence based, i.e., determined by the individual physician's expertise and expectation of treatment success⁷, combined with the best available clinical evidence from systematic research²⁹. Patient education on the choice of treatment and the expected outcomes seems to be a key factor to success and clinical negotiation appears to be an essential tool when it comes to discussing pain as an active

process involving patient expectations and coping strategies^{25,35}.

Patients with acute low back pain (less than 4 weeks of symptoms and without signs of specific LBP) can be treated with adequate pain medication and are recommended to keep as active as possible and resume work within 2 weeks. The treatment of patients with nonspecific LBP of 4 to 12 weeks has been called the “window of opportunity”²³. In these patients it is essential to prevent the onset of CLBP with long-term disability. From pain research it is known that the processing of pain occurs at three levels (peripheral, spinal and the brain) and the longer pain in an organ or a nerve lasts, the more chance that the perception of the pain is altered and that it becomes chronic, regardless of the state of that specific organ or nerve. This phenomenon is called spinal sensitization and central nervous system plasticity¹¹. Although psychological factors alone rarely cause pain, they do trigger or exacerbate a pain episode and contribute to the distress and disability experienced by the patient. Therefore, in CLBP, treatment that aims at pain modulation at all three levels (peripheral, spinal and the brain) generally achieves better pain management than therapy that only targets on one site. Multidisciplinary therapy should provide a treatment regimen of evidence-based interventions in a stepwise active approach¹⁹, starting with intensive exercise programs or cognitive behavioral interventions that may be combined with pain intervention (medication or neural ablative procedures) or psychotherapy.

What could be the role of spinal fusion in this multidisciplinary approach?

In the 2006 European guidelines it was stated that “only if a prolonged period of acknowledged conservative treatment modalities fails to relieve symptoms, fusion surgery for CLBP may be considered”². However, we found that fusion results remain unpredictable and current prognostic tests are not able to improve surgical outcome. For patient selection, proper diagnostics that include the assessment of the individual patient’s potential to benefit from surgery, are prerequisite in order to conceive realistic predictions of treatment outcome. So, if we believe there is a role for spinal fusion in the treatment of CLBP, more high quality trials to determine prognostic accuracy of tests, as well as the treatment’s safety and its effect on pain, function and return to work will be needed¹². RCTs provide a high level of evidence in comparing effectiveness of two treatment modalities, but may be restricted by a

limited external validity, which means that the results cannot be generalized for all CLBP patients²⁸. Therefore, observational data with indeed a lower level of evidence but a higher feasibility of being closer to daily practice, should be collected as well. The implementation of currently developing well-organized spine registries^{27,31} should enable the spinal community to acquire such representative data of high validity. As clinicians, we need such data, not only to counsel our patients, but also to inform other health care providers, insurers and policy makers whether spinal fusion may truly be an effective treatment strategy for a particular subgroup of CLBP patients, and whether it can be embedded in the current, widely supported multidisciplinary active approach of CLBP.

CONCLUSION

Surgical indications for CLBP remain ill defined. The relationship between pain and a degenerative intervertebral disc is unclear, and at the current state of knowledge, lumbar spinal fusion for CLBP is no better or worse than cognitive behavioral therapy or intensive exercise programs. However, fusion has a higher complication rate and is more costly for society. Treatment of CLBP patients should proceed in a multidisciplinary stepwise fashion, from the least invasive to the most invasive treatment. Patient education and counseling on the pros and cons of all treatment options are mandatory. If we believe there is a future role for spinal fusion in the treatment of these patients, better patient selection with subsequent improved outcome compared to conservative programs should be realized. For now, however, tests that are used in daily practice to predict the outcome of lumbar fusion do not succeed in identifying a subgroup of patients that will reliably benefit from fusion. Therefore, as its results remain unpredictable, spinal fusion is not to be recommended as a standard treatment for CLBP. The main message should be to encourage patients to take an active course of treatment in order to prevent long-term disability and chronicity.

REFERENCES

1. Abenhaim L, Rossignol M, Valat JP, et al. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. *Spine (Phila Pa 1976)* 2000;25:1S-33S.
2. Airaksinen O, Brox JJ, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006;15 Suppl 2:S192-300.
3. Aprill C, Bogduk N. High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. *Br J Radiol* 1992;65:361-9.
4. Blumenthal S, McAfee PC, Guyer RD, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. *Spine* 2005;30:1565-75; discussion E387-91.
5. Boden SD, McCowin PR, Davis DO, et al. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:1178-84.
6. Brox JJ, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913-21.
7. Campbell C, Guy A. 'Why can't they do anything for a simple back problem?' A qualitative examination of expectations for low back pain treatment and outcome. *J Health Psychol* 2007;12:641-52.
8. Carragee EJ. Clinical practice. Persistent low back pain. *N Engl J Med* 2005;352:1891-8.
9. Carragee EJ, Tanner CM, Khurana S, et al. The rates of false-positive lumbar discography in select patients without low back symptoms. *Spine* 2000;25:1373-80; discussion 81.
10. de Kleuver M, Oner FC, Jacobs WC. Total disc replacement for chronic low back pain: background and a systematic review of the literature. *Eur Spine J* 2003;12:108-16.
11. DeLeo JA. Basic science of pain. *J Bone Joint Surg Am* 2006;88 Suppl 2:58-62.
12. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery - the case for restraint. *N Engl J Med* 2004;350:722-6.
13. Fairbank J, Frost H, Wilson-MacDonald J, et al. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *Bmj* 2005;330:1233.

14. Freeman BJ, Davenport J. Total disc replacement in the lumbar spine: a systematic review of the literature. *Eur Spine J* 2006;15 Suppl 3:S439-47.
15. Fritzell P, Hagg O, Jonsson D, et al. Cost-effectiveness of lumbar fusion and nonsurgical treatment for chronic low back pain in the Swedish Lumbar Spine Study: a multicenter, randomized, controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2004;29:421-34; discussion Z3.
16. Fritzell P, Hagg O, Wessberg P, et al. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001;26:2521-32; discussion 32-4.
17. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 1999;24:1820-32.
18. Guyer RD, McAfee PC, Banco RJ, et al. Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: five-year follow-up. *Spine J* 2009;9:374-86.
19. Guzman J, Esmail R, Karjalainen K, et al. Multidisciplinary rehabilitation for chronic low back pain: systematic review. *Bmj* 2001;322:1511-6.
20. Haldeman S, Dagenais S. A supermarket approach to the evidence-informed management of chronic low back pain. *Spine J* 2008;8:1-7.
21. Ito M, Incorvaia KM, Yu SF, et al. Predictive signs of discogenic lumbar pain on magnetic resonance imaging with discography correlation. *Spine (Phila Pa 1976)* 1998;23:1252-8; discussion 9-60.
22. Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 1994;331:69-73.
23. Nordin M, Balague F, Cedraschi C. Nonspecific lower-back pain: surgical versus nonsurgical treatment. *Clin Orthop Relat Res* 2006;443:156-67.
24. Ostelo RW, van Tulder MW, Vlaeyen JW, et al. Behavioural treatment for chronic low-back pain. *Cochrane Database Syst Rev* 2005:CD002014.
25. Risdon A, Eccleston C, Crombez G, et al. How can we learn to live with pain? A Q-methodological analysis of the diverse understandings of acceptance of chronic pain. *Soc Sci Med* 2003;56:375-86.

26. Rivero-Arias O, Campbell H, Gray A, et al. Surgical stabilisation of the spine compared with a programme of intensive rehabilitation for the management of patients with chronic low back pain: cost utility analysis based on a randomised controlled trial. *Bmj* 2005;330:1239.
27. Roder C, Muller U, Aebi M. The rationale for a spine registry. *Eur Spine J* 2006;15 Suppl 1:S52-6.
28. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" *Lancet* 2005;365:82-93.
29. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *Bmj* 1996;312:71-2.
30. Staal JB, Hlobil H, Twisk JW, et al. Graded activity for low back pain in occupational health care: a randomized, controlled trial. *Ann Intern Med* 2004;140:77-84.
31. Stromqvist B, Jonsson B, Fritzell P, et al. The Swedish National Register for lumbar spine surgery: Swedish Society for Spinal Surgery. *Acta Orthop Scand* 2001;72:99-106.
32. van den Hout JH, Vlaeyen JW, Heuts PH, et al. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. *Clin J Pain* 2003;19:87-96.
33. van Ooij A, Oner FC, Verbout AJ. Complications of artificial disc replacement: a report of 27 patients with the SB Charite disc. *J Spinal Disord Tech* 2003;16:369-83.
34. van Tulder M, Malmivaara A, Esmail R, et al. Exercise therapy for low back pain: a systematic review within the framework of the cochrane collaboration back review group. *Spine (Phila Pa 1976)* 2000;25:2784-96.
35. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain* 2000;85:317-32.
36. Waddell G. Low back pain: a twentieth century health care enigma. *Spine* 1996;21:2820-5.
37. Weishaupt D, Zanetti M, Hodler J, et al. Painful Lumbar Disk Derangement: Relevance of Endplate Abnormalities at MR Imaging. *Radiology* 2001;218:420-7.
38. Wilson-MacDonald J, Fairbank J, Frost H, et al. The MRC spine stabilization trial: surgical methods, outcomes, costs, and complications of surgical stabilization. *Spine* 2008;33:2334-40.

SUMMARY

The general introduction in **chapter 1** posits the problem of chronic low back pain (CLBP), one of the main causes of disability in the western world with a huge economic burden to society. For CLBP no specific underlying anatomic cause has been identified. Imaging often reveals degenerative findings of the disc or facet joints of one or more lumbar motion segments. These findings, however, can also be observed in asymptomatic people. It has been suggested that pain in degenerated discs may be caused by the ingrowth of nerve fibers into tears or clefts of the annulus fibrosus or nucleus pulposus, and by reported high levels of pro-inflammatory mediators. As this so-called discogenic pain is often exacerbated by mechanical loading, the concept of relieving pain by spinal fusion to stabilize a painful spinal segment, has been developed. For some patients lumbar spinal fusion indeed is beneficial, but its results are highly variable and hard to predict for the individual patient. To identify those CLBP patients who will benefit from fusion, many surgeons rely on tests that are assumed to predict the outcome of spinal fusion. The three most commonly used prognostic tests in daily practice are immobilization in a lumbosacral orthosis, provocative discography, and trial immobilization by temporary external transpedicular fixation. Aiming for consensus on the indications for lumbar fusion and in order to improve its results by better patient selection, it is essential to know the role and value of these prognostic tests for CLBP patients in clinical practice.

In **chapter 2**, the results of a national survey among spine surgeons in the Netherlands are presented. The surgeons were questioned about their opinion on prognostic factors and about the use of predictive tests for lumbar fusion in CLBP patients. The comments were compared with findings from the prevailing literature. The survey revealed a considerable lack of uniformity in the use and appreciation of predictive tests. Prognostic factors known from the literature were not consistently incorporated in the surgeons' decision making process either. This heterogeneity in strategy is most probably caused by the lack of sound scientific evidence for current predictive tests and it was concluded that currently there is not enough consensus among spine surgeons in the Netherlands to create national guidelines for surgical decision making in CLBP.

In **chapter 3a**, the hypothesized working mechanism of a pantaloons cast (i.e., minimization of lumbosacral joint mobility) was studied. In patients who were admitted for a temporary external transpedicular fixation test (TETF), infrared light markers were fixed to the protruding pins of two spinal levels. In this way three-dimensional motion between these levels could be analyzed opto-electronically. During dynamic test conditions such as walking, a plaster cast, either with or without unilateral hip fixation, did not significantly decrease lumbosacral joint motion.

Although not substantiated by sound scientific support, lumbosacral orthoses or pantaloons casts are often used in everyday practice as a predictor for the outcome of fusion. In **chapter 3b**, a systematic review of the literature supplemented with a prospective cohort study was performed in order to assess the value of a pantaloons cast in surgical decision-making. It appeared that only in chronic LBP patients with no prior spine surgery, a pantaloons cast test with substantial pain relief suggests a favorable outcome of lumbar fusion compared to conservative treatment. In patients with prior spine surgery the test was of no value.

It is believed by many spine surgeons that provocative discography, unlike plain radiographs or magnetic resonance imaging, is a physiologic test that can truly determine whether a disc is painful and relevant in a patient's pain syndrome, irrespective of the morphology of the disc. It has been suggested that in order to achieve a successful clinical outcome of lumbar fusion, suspect discs should be painful and adjacent control discs should elicit no pain on provocative discography. In **chapter 4a**, a cohort of patients in whom the decision to perform lumbar fusion was based on an external fixation (TETF) trial, was studied. The results of preoperative discography of solely the levels adjacent to the fusion were analyzed retrospectively. It appeared that in this select group of patients the discographic status of discs adjacent to a lumbar fusion did not have any effect on the clinical outcome.

The most feared complication of lumbar discography is discitis. Although low in incidence, this is a serious complication for a diagnostic procedure and prevention by the use of prophylactic antibiotics has been advocated. In search for clinical

guidelines, the risk of postdiscography discitis was studied in **chapter 4b**, by means of a systematic literature review and a cohort of 200 consecutive patients. Without the use of prophylactic antibiotics, an overall incidence of postdiscography discitis of 0.25% was found. To prove that antibiotics would actually prevent discitis, a randomized trial of 9,000 patients would be needed to reach significance. Given the possible adverse effects of antibiotics, it was concluded that the routine use of prophylactic antibiotics in lumbar discography is not indicated.

In **chapter 5**, the middle- and long-term results of external fixation (TETF) as a test to predict the clinical outcome of lumbar fusion were studied in a group of back pain patients for whom there was doubt about the indication for surgery. The test included a placebo trial, in which the patient was unaware whether the lumbar segmental levels were fixed or dynamized. Using strict and objective criteria of pain reduction on a visual analogue scale, the TETF test failed to predict clinical outcome of fusion in this select group of patients. Pin track infection and nerve root irritation were registered as complications of this invasive test. It was concluded that in chronic low back pain patients with a doubtful indication for fusion, TETF is not recommended as a supplemental tool for surgical decision-making.

Chapter 6 presents the results of a systematic literature review concerning the prognostic accuracy of tests that are currently used in clinical practice and that are presumed to predict the outcome of lumbar spinal fusion for CLBP. The tests of interest were magnetic resonance imaging (MRI), TLSO immobilization, TETF, provocative discography and facet joint infiltration. Only 10 studies reporting on three different index tests (discography, TLSO immobilization and TETF) could be selected. It appeared that the overall methodological quality of included studies was poor and the determined accuracy of the prognostic tests was low. Higher quality trials that include negatively tested as well as positively tested patients for fusion are needed in order to be able to calculate test qualifiers, such as sensitivity, specificity and likelihood ratios, to truly determine the prognostic accuracy of these tests. It was concluded that best evidence does not support the use of any prognostic test in clinical practice. At present, no subset of patients with low-back pain can be identified, for whom spinal fusion is a reliable and effective treatment.

The general discussion in **chapter 7** focuses on the current and possible future role of lumbar spinal fusion for CLBP, and is based on work from the present thesis combined with findings from the literature. Causality of nonspecific spinal pain is complex and CLBP should be regarded as a symptom rather than a diagnosis in patients with different stages of impairment and disability. The so-called biopsychosocial model for the evaluation of CLBP patients is explained. This model aims at identifying underlying psychosocial factors as well as biological factors. Patient triage in a multidisciplinary setting or Spine Centre is advocated and the value of other treatment options, such as cognitive-behavioral therapy and intensive exercise programs, is further elaborated. Treatment should occur in a stepwise fashion starting with the least invasive treatment. The current, widely supported active approach of CLBP is recommended, and emphasis should be laid on self-management and empowerment of patients to take an active course of treatment in order to prevent long-term disability and chronicity.

SAMENVATTING

In de algemene introductie van **hoofdstuk 1** wordt het probleem chronische lage rugpijn, één van de belangrijkste oorzaken van arbeidsongeschiktheid in de Westerse wereld met hoge kosten voor de maatschappij, uiteengezet. Als oorzaak voor chronische lage rugpijn is tot op heden geen specifieke onderliggende anatomische afwijking gevonden. Beeldvormende diagnostiek toont vaak tekenen van degeneratie van de discus intervertebralis en eventueel de facetgewrichten van één of meerdere lumbale bewegingssegmenten. Deze bevindingen worden echter ook vaak gezien bij mensen zonder rugklachten. Pijn in een gedegenererde discus wordt mogelijk veroorzaakt door het ontstaan van scheuren in de anulus fibrosus of nucleus pulposus tijdens degeneratie waarbij ingroei van sensibele zenuwvezels is gevonden. Tevens zijn hoge concentraties van pro-inflammatoire mediators gerapporteerd. Aangezien deze zogenaamde discogene pijn vaak verergerd wordt door mechanisch belasten, is de gedachte ontstaan om pijn te verlichten middels het vastzetten van een verondersteld pijnlijk bewegingssegment, oftewel spondylodese. Dit geeft inderdaad voor een aantal patiënten adequate pijnvermindering, maar de resultaten zijn sterk wisselend en moeilijk te voorspellen voor de individuele patiënt.

Om patiënten die baat zullen hebben van een spondylodese te selecteren, vertrouwen veel wervelkolomspecialisten op tests die verondersteld worden het resultaat van spondylodese te kunnen voorspellen. De 3 meest gebruikte prognostische testen in de dagelijkse praktijk zijn proef-immobilisatie in een lumbosacrale orthose, provocatie discografie, en test-immobilisatie middels tijdelijke externe fixatie.

Om consensus te creëren betreffende de indicatie-stelling tot spondylodese voor chronische lage rugpijn, alsmede om de resultaten hiervan te kunnen verbeteren door betere patiëntselectie, is het essentieel om de rol en waarde van bovengenoemde prognostische tests voor de patiënt in de dagelijkse praktijk te kennen.

In **hoofdstuk 2** worden de resultaten van een nationale enquête onder Nederlandse wervelkolomchirurgen gepresenteerd. De chirurgen werd gevraagd naar hun mening over prognostische factoren, alsmede naar hun gebruik van tests

die het klinische resultaat van spondylodese voor chronische lage rugpijn kunnen voorspellen. De antwoorden werden vergeleken met de heersende opvattingen uit de vakliteratuur. De enquête toonde een gebrek aan uniformiteit in gebruik en waardering van de voorspellende tests. Prognostische factoren bekend uit de literatuur werden niet consequent betrokken in de besluitvorming tot spondylodese. Deze heterogeniteit in behandelingsstrategie wordt hoogstwaarschijnlijk veroorzaakt door het gebrek aan wetenschappelijk bewijs voor de waarde van deze voorspellende tests. Er werd geconcludeerd dat er momenteel onvoldoende consensus onder Nederlandse wervelkolomchirurgen bestaat om nationale richtlijnen voor de indicatiestelling tot spondylodese bij chronische lage rugpijnpatiënten op te kunnen stellen.

De intentie van een proeftest met een gipskorset met pijp is het minimaliseren van lumbosacrale mobiliteit om zo het effect van een spondylodese te simuleren. Dit effect werd in **hoofdstuk 3a** bestudeerd in een groep patiënten met chronische lage rugpijn bij wie twijfel over de indicatie tot spondylodese bestond en die opgenomen waren voor een tijdelijke transpediculaire externe fixatietest (zie hoofdstuk 7). Hierbij werden 2 metalen frames met infrarood licht reflecterende bollen rigide gefixeerd op de percutaan uitstekende uiteinden van transpediculaire schroeven in 2 wervelsegmenten. Zo kon het driedimensionale bewegingspatroon tussen deze segmenten opto-elektronisch bepaald worden. Het bleek dat onder dynamische test condities, zoals tijdens lopen, een gipskorset met pijp geen significante vermindering van lumbosacrale mobiliteit bewerkstelligt.

Hoewel er geen gedegen wetenschappelijke onderbouwing voor is, worden thoracolumbosacrale orthoses (TLSO) of gipskorsetten in de dagelijkse praktijk regelmatig gebruikt om het klinische resultaat van een spondylodese te voorspellen. In **hoofdstuk 3b** wordt de waarde van proefimmobilisatie met een TLSO bestudeerd in een patiënten cohort, alsmede middels een systematische review van de literatuur. Alleen bij patiënten zonder eerdere rugoperatie bleek dat, indien zij substantiële pijnvermindering tijdens het dragen van de TLSO ondervonden, het klinische resultaat van een spondylodese significant beter was dan van conventionele conservatieve behandeling. Voor patiënten die eerder wegens

rugklachten waren geopereerd (zogenaamde failed back surgery patiënten) bleek de test geen voorspellende waarde te hebben.

Voorstanders van discografie stellen dat, in tegenstelling tot röntgenfoto's of een MRI-scan, provocatie middels discografie een fysiologische test is waarmee bij een patiënt bepaald kan worden of een discus, ongeacht morfologische kenmerken, als oorzaak van de pijn beschouwd mag worden. Zo wordt gesuggereerd dat om een goed klinisch resultaat van spondylodese te mogen verwachten, discografie van de te fixeren discus duidelijk herkenbare pijn moet opwekken, terwijl bij aangrenzende "controle" disci geen herkenbare pijn geprovoceerd mag worden. In **hoofdstuk 4a** werd een cohort van patiënten bestudeerd die provocatie discografie hadden ondergaan, maar waarbij de indicatie tot spondylodese was gesteld op basis van een tijdelijke transpediculaire externe fixatietest (zie hoofdstuk 7). De resultaten van preoperatieve provocatie middels discografie van de aangrenzende "controle" disci werden retrospectief geanalyseerd. Bij deze patiënten bleek het latere klinische resultaat van spondylodese niet bepaald te worden door de preoperatieve discografische status (wel of geen herkenbare pijn bij provocatie middels discografie) van aangrenzende "controle" disci.

Hoewel de incidentie laag is, wordt discitis als de meest gevreesde complicatie van discografie beschouwd. Aangezien het een relatief ernstige complicatie voor een diagnostische procedure betreft, wordt veelvuldig het routinematig gebruik van antibiotica profylaxe bij discografie geadviseerd. Op zoek naar standaardisering van klinische richtlijnen werd het risico op discitis na discografie bestudeerd in **hoofdstuk 4b**. In een systematische review van de literatuur, aangevuld met een cohortstudie van 200 opeenvolgende patiënten waarbij discografie zonder antibiotica profylaxe werd uitgevoerd, werd een discitis incidentie van 0,25% gevonden. Om te kunnen bewijzen dat antibiotica profylaxe daadwerkelijk het risico op postdiscografie discitis significant kan verminderen, zou een randomised trial (RCT) van 9.000 patiënten nodig zijn. Gezien de mogelijke bijwerkingen van antibiotica en risico op resistentie-vorming, werd geconcludeerd dat routinematige antibiotica profylaxe bij lumbale discografie niet is geïndiceerd. Uiteraard dient de procedure onder strikt steriele omstandigheden verricht te worden.

In **hoofdstuk 5** wordt de waarde bestudeerd van tijdelijke externe transpediculaire fixatie (TETF) als test om het lange termijn klinische resultaat van spondylodese te voorspellen in een cohort patiënten waarbij twijfel over de indicatie tot spondylodese bestond. Bij de test was ook de zogenaamde “placebo-stand” inbegrepen, waarbij zonder dat de patiënt het wist, de fixerende staven horizontaal werden geplaatst zodat er geen daadwerkelijke immobilisatie van 2 wervelniveaus plaatsvond. De mate van pijnvermindering die de patiënt ondervond tijdens de test werd genoteerd op een visueel analoge schaal (VAS) waarbij 30% pijnreductie als klinisch relevant werd beschouwd. De test bleek bij deze selecte patiëntengroep niet in staat het klinische resultaat van spondylodese betrouwbaar te voorspellen. Schroefgat infectie en zenuwwortel irritatie werden frequent geregistreerd als complicatie van deze invasieve test. Geconcludeerd werd dat bij lage rugpatiënten waarbij een onzekere indicatie tot spondylodese bestaat na routine diagnostiek, TETF niet aanbevolen dient te worden als extra test voor klinische besluitvorming.

Hoofdstuk 6 presenteert een systematische review van de literatuur betreffende de validiteit van prognostische tests die momenteel in de dagelijkse praktijk gebruikt worden om het klinische resultaat van spondylodese bij chronische lage rugpijn te voorspellen. Onderzocht werden: MRI, corset-immobilisatie, tijdelijke externe fixatie, provocatie discografie en facetgewricht blokkade. Slechts 10 studies waarbij voorspellende indicatoren als sensitiviteit en specificiteit waren bepaald, konden geïnccludeerd worden. De prognostische accuratesse van de geïdentificeerde tests was laag. Bovendien was de methodologische kwaliteit van vrijwel alle geselecteerde studies matig tot slecht zodat geen definitieve uitspraken gedaan konden worden. Om de daadwerkelijke prognostische waarde van bovengenoemde tests te kunnen bepalen, zullen studies van betere kwaliteit qua opzet en uitvoering verricht dienen te worden. Hierbij dienen ook patiënten met een negatieve testuitslag geïnccludeerd te worden om veelvuldig geconstateerde verificatie bias te voorkomen. Geconcludeerd werd dat er geen wetenschappelijk bewijs is dat één van bovengenoemde prognostische tests van waarde is voor de indicatiestelling tot spondylodese in de dagelijkse praktijk. Momenteel kan bij patiënten met lage rugpijn niet betrouwbaar voorspeld worden of ze gebaat zullen zijn met een spondylodese.

De algemene discussie in **hoofdstuk 7** behandelt de huidige en mogelijk toekomstige rol van lumbale spondylodese als behandeling van chronische lage rugpijn, en is gebaseerd op werk uit de huidige thesis gecombineerd met bevindingen uit de literatuur. Lage rugpijn zonder specifieke diagnose is vaak multicausaal bepaald en dient meer als een symptoom in patiënten met een bepaalde mate van beperking of invaliditeit beschouwd te worden, dan als diagnose.

Het zogenaamde biopsychosociale model ter evaluatie van lage rugpijnpatiënten wordt toegelicht. Dit model streeft ernaar om onderliggende psychosociale problematiek bij patiënten met fysieke klachten te identificeren en zo nodig (mede) te behandelen. In dat kader wordt triage van rugpatiënten in een multidisciplinaire setting of Spine Centrum aanbevolen. Bovendien wordt de waarde van conservatieve behandelingen als cognitieve gedragstherapie en gestructureerde intensieve oefenprogramma's onderstreept. Behandeling van chronische lage rugpijn dient stapsgewijs uitgevoerd te worden, te beginnen met de minst invasieve therapie. Cruciaal hierbij is patiënten te motiveren hun probleem aan te pakken met een attitude van self management en een actieve levensstijl met zo nodig oefenprogramma's onder begeleiding. Doel van deze in toenemende mate gepropageerde, actieve benadering van lage rugpijn is het eindstadium van chroniciteit, en de daarmee gepaard gaande invaliditeit en arbeidsongeschiktheid, te voorkomen.

DANKWOORD

Bij deze wil ik allen bedanken die een bijdrage hebben geleverd aan mijn proefschrift.

Allereerst gaat mijn dank uit naar mijn promotoren Geert Walenkamp en Rob de Bie voor alle regelwerkzaamheden op het laatste moment, en naar mijn co-promotor Marinus de Kleuver voor zijn deskundige en immer enthousiasmerende begeleiding.

In het bijzonder wil ik Patsy Anderson, Wilco Jacobs, Bart Nienhuis en Bart Staal hartelijk danken voor hun hulp bij methodologie en analyse van de diverse studies. Hun professionele ondersteuning was onontbeerlijk voor de totstandkoming van dit proefschrift.

CURRICULUM VITAE PAUL WILLEMS

Born in Sint-Anthonis, The Netherlands, on October 14, 1965

1978-1984: Secondary School, Atheneum, Boxmeer

1984-1988: Medical Degree, Katholieke Universiteit Nijmegen

1988-1991: Medical Doctor: Internships of Medicine including Fellowship Primary Health Care, Nicaragua, Central America (1990)

1992-1993: Orthopaedic research fellow, Sint Maartenskliniek Nijmegen and Academisch Ziekenhuis Nijmegen

1994-1995: Resident in General Surgery, Lucas Ziekenhuis, Apeldoorn

1995-1996: Resident in Orthopaedics, St. Elisabeth Hospital, Willemstad, Curaçao, Netherlands Antilles

1996-1998: Training in General Surgery, Atrium Medisch Centrum, Heerlen

1998-2002: Training in Orthopaedic Surgery at the Atrium Medisch Centrum, Heerlen and Academisch Ziekenhuis Maastricht

2002-2003: AO-Fellowship in Spine Surgery, Sint Maartenskliniek Nijmegen

2004-2007: Orthopaedic surgeon, Viecuri Medisch Centrum, Venlo

01-09-2007: Orthopaedic surgeon, subspecialty Spine, Maastricht Universitair Medisch Centrum, The Netherlands.

